Quantum Computing without Magic

Scientific and Engineering Computation
William Gropp and Ewing Lusk, editors; Janusz Kowalik, founding editor

Data-Parallel Programming on MIMD Computers, Philip J. Hatcher and Michael J. Quinn, 1991

Unstructured Scientific Computation on Scalable Multiprocessors, edited by Piyush Mehrotra, Joel Saltz, and Robert Voigt, 1992

Parallel Computational Fluid Dynamics: Implementation and Results, edited by Horst D. Simon, 1992

Enterprise Integration Modeling: Proceedings of the First International Conference, edited by Charles J. Petrie, Jr., 1992

The High Performance Fortran Handbook, Charles H. Koelbel, David B. Loveman, Robert S. Schreiber, Guy L. Steele Jr. and Mary E. Zosel, 1994

PVM: Parallel Virtual Machine–A Users' Guide and Tutorial for Network Parallel Computing, Al Geist, Adam Beguelin, Jack Dongarra, Weicheng Jiang, Bob Manchek, and Vaidy Sunderam, 1994

Practical Parallel Programming, Gregory V. Wilson, 1995

Enabling Technologies for Petaflops Computing, Thomas Sterling, Paul Messina, and Paul H. Smith, 1995

An Introduction to High-Performance Scientific Computing, Lloyd D. Fosdick, Elizabeth R. Jessup, Carolyn J. C. Schauble, and Gitta Domik, 1995

Parallel Programming Using C++, edited by Gregory V. Wilson and Paul Lu, 1996

Using PLAPACK: Parallel Linear Algebra Package, Robert A. van de Geijn, 1997

Fortran 95 Handbook, Jeanne C. Adams, Walter S. Brainerd, Jeanne T. Martin, Brian T. Smith, Jerrold L. Wagener, 1997

MPI—The Complete Reference: Volume 1, The MPI Core, Marc Snir, Steve Otto, Steven Huss-Lederman, David Walker, and Jack Dongarra, 1998

MPI—The Complete Reference: Volume 2, The MPI-2 Extensions, William Gropp, Steven Huss-Lederman, Andrew Lumsdaine, Ewing Lusk, Bill Nitzberg, William Saphir, and Marc Snir, 1998

A Programmer's Guide to ZPL, Lawrence Snyder, 1999

How to Build a Beowulf, Thomas L. Sterling, John Salmon, Donald J. Becker, and Daniel F. Savarese, 1999

Using MPI: Portable Parallel Programming with the Message-Passing Interface, second edition, William Gropp, Ewing Lusk, and Anthony Skjellum, 1999

Using MPI-2: Advanced Features of the Message-Passing Interface, William Gropp, Ewing Lusk, and Rajeev Thakur, 1999

Beowulf Cluster Computing with Linux, Thomas Sterling, 2001

Beowulf Cluster Computing with Windows, Thomas Sterling, 2001

Scalable Input/Output: Achieving System Balance, Daniel A. Reed, 2003

Quantum Computing without Magic

Draft of October 1, 2006

Zdzisław Meglicki

Dedicated to my cats, Bambosz and Sofa, who have a bone to pick with Dr
Schrödinger.

Contents

Series Foreword

A book in a series usually has a series foreword, written by the series editor.

William Gropp and Ewing Lusk, Editors
Janusz Kowalik, Founding Editor

Foreword

This is the foreword for the book, sometimes written by a well-known figure in the field.

Able Baker, President
HAL International
Urbana, IL, USA
July 4th, 1999

Preface

 The preface can give background for the book, including some history of the field and more personal notes.

Acknowledgments

Here you thank everyone who helped you put your book together.

D. E. D. Tors
H. S. Elper

Quantum Computing without Magic

"If you want to amount to anything as a witch, Magrat Garlick, you got to learn three things. What's real, what's not real, and what's the difference—"

Terry Pratchett, Witches Abroad

1 Bits and Registers

1.1   Physical embodiments of a bit

Information technology devices, such as desktop computers, laptops, palmtops, *bits and bytes*
cellular phones, DVD players and many others have pervaded our every day life to
such extent that it is difficult to find a person who would not have at least some
idea about what bits and bytes are. I shall assume therefore that the reader knows
about both, enough to understand that a bit is the "smallest indivisible chunk of
information" and that a byte is a string of eight bits.

   Yet the concept of a bit as the "smallest indivisible chunk of information" is a *discretization of*
somewhat stifling convention. It is possible to dose information in any quantity, *information is a*
not necessarily in discrete chunks, and this is how many analogue devices, including *convention*
analogue computers, work. What more, it takes a considerable amount of signal
processing, and consequently also power and time, to maintain a nice rectangular
shape of pulses representing bits in digital circuits. Electronic circuits that can
handle information directly, without chopping it to bits and arranging it into bytes,
can be orders of magnitude faster and more energy efficient than digital circuits.

   How are bits and bytes actually stored, moved and processed inside digital de- *storing and*
vices? There are many ways to do this. Figure 1.1 shows a logic diagram of one of *manipulating*
the simplest memory cells, a flip-flop. *bits*
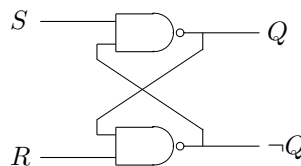
Figure 1.1: A very simple flip-flop comprising two cross-coupled NAND gates.

   The flip-flop in the figure comprises two cross-coupled NAND gates. It is easy to *a flip-flop can*
analyze the behavior of the circuit. Suppose R is set to 0 and S is set to 1. If R *be used to*
is 0 then regardless of what the second input to the NAND gate at the bottom is, *implement a*
its output must be 1. Therefore the second input to the NAND gate at the top is *1-bit memory*
1 and so its output Q must be 0. Because the roles of R and S in the device are *cell*

completely symmetric, this result implies that if R is set to 1 and S to 0 we'll get that $Q = 1$ and $\neg Q = 0$. Table 1.1 sums up these simple results.

R	S	Q	$\neg Q$
0	1	0	1
1	0	1	0

Table 1.1: Q and $\neg Q$ as functions of R and S for the flip-flop of Figure 1.1.

Observe that once the value of Q has been set to either 0 or 1, setting both R and S to 1 retains the pre-set value of Q. This is easy to see. Suppose Q has been set to 1. Therefore $\neg Q$ is 0 and so one of the inputs to the upper NAND gate is 0, which implies that its output must be 1 indeed. In order for $\neg Q$ to be 0, both inputs to the lower NAND gate must be 1, and so they are because $R = 1$.

Now suppose that Q has been pre-set to 0 instead. Hence the second input to the lower NAND gate is 0 and therefore the output of the gate, $\neg Q$ is 1, which is exactly what is required in order for Q to be 1, on account of $\neg Q$ being the second input to the upper NAND gate.

And so our flip-flop behaves like a simple memory device. By operating on its inputs we can set its output to either 0 or 1 and then by setting both inputs to 1 we can make it remember the pre-set state.

what is inside   It is instructive to have a closer look at what happens inside the NAND gates when
the NAND *gate*   the device remembers its pre-set state. How is this remembering accomplished?
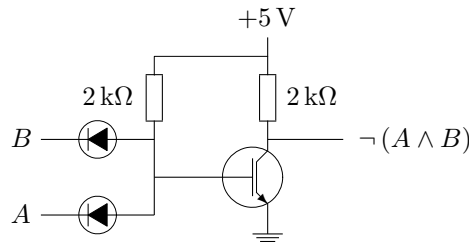
Figure 1.2: A diode-transistor-logic implementation of a NAND gate.

Figure 1.2 shows a simple diode-transistor logic (DTL) implementation of a NAND gate. Each of the diodes on the two input lines A and B conducts when 0 is applied to its corresponding input. The diodes disconnect when 1 is applied to their inputs. The single transistor in the circuit is an n-channel transistor. This means that the

channel of the transistor conducts when a positive charge, logical 1, is applied to the gate. Otherwise the channel blocks. Let us consider what is going to happen if any of the two inputs is set to 0. In this case the corresponding diode conducts and the positive charge drains from the gate of the transistor. Consequently its channel blocks and the output of the circuit ends up being 1. On the other hand if both inputs are set to 1, both diodes block. In this case positive charge flows towards the gate of the transistor and accumulates there and the transistor channel conducts. This sets the potential on the output line to 0. The resulting truth table of the device is shown in table 1.2. This is indeed the table of a NAND gate.

A	B	$\neg(A \wedge B)$
0	0	1
0	1	1
1	0	1
1	1	0

Table 1.2: The truth table of the DTL NAND gate shown in figure 1.2.

The important thing to observe in the context of our considerations is that it is the presence or the absence of the charge on the transistor gate that determines the value of the output line. If there is no accumulation of positive charge on the gate, the output line is set to 0, if there is a sufficient positive charge on the gate, the output line is set to 1.

Returning to our flip-flop example, we can now see that the physical embodiment of the bit, which the flip-flop "remembers", is the electric charge stored on the gate of the transistor located in the upper NAND gate of the flip-flop circuit. If there is an accumulation of positive charge on the transistor's gate, the Q line of the flip-flop becomes 0 and if the charge has drained from the gate, the Q line becomes 1. The Q line itself merely provides us with the means of reading the bit. *the charge stored on the gate of a transistor represents the bit*

We could replace the flip-flop simply with a box and a pebble. An empty box would correspond to a drained transistor gate, and this we could then *read* as 1, and if we find a pebble in the box, we would read this as 0. The box and the pebble would work very much like the flip-flop in this context.

It is convenient to reverse the convention and read a pebble in the box as 1 and its absence as 0. We could do the same, of course, with the flip-flop, simply by renaming Q to $\neg Q$ and vice versa.

Seemingly we have performed an act of conceptual digitization in discussing and then translating the physics of the flip-flop and of the DTL NAND gate to *the box and the pebble* picture.

the two states of a transistor are connected by a continuous transition

A transistor is really an analog amplifier and it is possible to apply any potential to its gate, which yields a range of continuous values to its channel's resistance. In order for the transistor to behave like a switch, and for the circuit presented in figure 1.2, to behave like a NAND gate, we must condition its input and output voltages: these are usually restricted to $\{0\,\mathrm{V}, +5\,\mathrm{V}\}$ and switched very rapidly between the two values. Additionally parts of the circuit may be biased at $-5\,\mathrm{V}$ in order to provide adequate polarization. Even then, when looked at with an oscilloscope, pulses representing bits do not have sharp edges. Rather there are transients there and these must be analyzed rigorously at the circuit design stage in order to eliminate unexpected faulty behavior.

On the other hand, the presence or the absence of the pebble in the box apparently represents two very distinct separate states. There are no transients here. The pebble either is or is not in the box. *Tertium non datur.*

the two states of a pebble in a box are connected by a continuous transition too many-valued logic

Yet, observe that even this is a convention, because, for example, we could place the pebble in such a way that only a half of it would be in the box, and the other half would be outside. How should we account for this situation?

In binary, digital logic we ignore such states. But there are other types of logic, which allow for the pebble to be half-way or a third or any other portion of it within the box. Such logic systems fall under the category of *many-valued logics* [22], some of which are even *infinitely-valued*. An example of an infinitely-valued logic is the popular *fuzzy logic* [38] commonly used in robotics, data bases, image processing and expert systems.

quantum logic as an infinitely-valued logic

When we get to look at quantum logic more closely these considerations will acquire a new deeper meaning, which will eventually lead to the notion of superposition of quantum states. Quantum logic is one of these systems, where a pebble can be half way in one box and half way in another one. And the boxes don't even have to be adjacent, sic!

1.2   Registers

a register

A row of flip-flops connected with each other in various ways constitutes a register. Depending on how the flip-flops are connected the register may be used just as a store, or it can be used to perform some arithmetic operations.

a 3-bit modulo-7 counter

Figure 1.3 shows a simple three-bit modulo-7 counter implemented with three JK flip-flops. A JK flip-flop is a more complex device than the one shown in
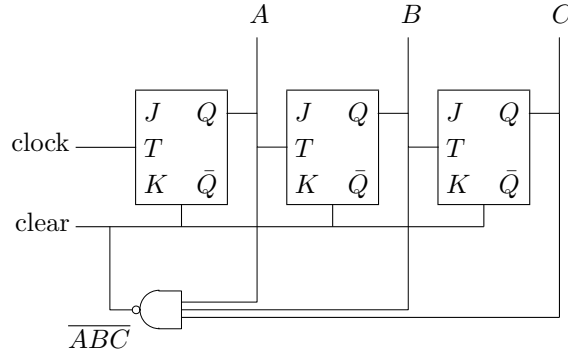
Figure 1.3: A modulo-7 counter made of three JK flip-flops.

figure 1.1, but all that the reader needs to know in order to understand how the counter works are the following two rules:

1. The state Q of the JK flip-flop toggles on the *trailing* edge of the clock pulse T, i.e., when the state of the input T changes from 1 to 0;

2. Applying 0 to *clear* resets Q to 0.

Assume that the whole counter starts in the $\{C = 0, B = 0, A = 0\}$ state. On the first application of the pulse to the *clock* input, A toggles to 1 on the trailing edge of the pulse and stays there. The state of the register becomes $\{C = 0, B = 0, A = 1\}$. On the second application of the clock pulse A toggles back to 0, but this change now toggles B to 1 and so the state of the register becomes $\{C = 0, B = 1, A = 0\}$. On the next trailing edge of the clock pulse A toggles to 1 and the state of the register is now $\{C = 0, B = 1, A = 1\}$. When A toggles back to 0 on the next application of the clock pulse, this triggers the change in B from 1 to 0, but this in turn toggles C and so the state of the register becomes $\{C = 1, B = 0, A = 0\}$, and so on. Dropping $C =$, $B =$ and $A =$ from our notation describing the state of the register we can see the following progression:

$$\{000\} \rightarrow \{001\} \rightarrow \{010\} \rightarrow \{011\} \rightarrow \{100\} \rightarrow \ldots$$

We can interpret the strings enclosed in curly brackets as binary numbers and upon having converted them to decimal notation we obtain

$$0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow \ldots$$

the states of the register

The device counts clock pulses. It does so by remembering the previous value and then adding 1 to it on detecting the trailing edge of the clock pulse. When A, B, and C all become 1 at the same time, the NAND gate at the bottom of the circuit applies *clear* to all three flip-flops and so A, B, and C get reset to 0. This happens so fast that the counter does not stay in the {111} configuration for an appreciable amount of time. It counts from 0 *through* 6 transitioning through seven distinct stable *states* in the process.

transitions between the states
     At first glance we may think that the counter *jumps* between the discrete states. A closer observation of the transitions with an oscilloscope shows that the counter *glides* between the states through a continuum of various configurations, which cannot be interpreted in terms of digital logic. But the configurations in the continuum are unstable and the gliding takes very little time, so that the notion of the *jumps* is a good approximation[1].

a register made of boxes and pebbles
     By now we know already that a possible physical embodiment of a bit is an accumulation of electric charge on the gate of a transistor inside a flip-flop. And we can also think about the presence or the absence of the charge on the gate in the same way we think about the presence or absence of a pebble in a box. And so, instead of working with a row of flip-flops we can work with a row of boxes and pebbles and such a system is also a register, albeit a much slower one and more difficult to manipulate too.

an "almost quantum" register
     The following figure shows an example of a box and a pebble register that displays some features that are reminiscent of quantum physics.
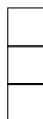
     The register contains 3 boxes, which are stacked vertically. Their position corresponds to the energy of a pebble that may be placed in a box. The higher the location of the box, the higher the energy of the pebble. The pebbles that are used in the register have a peculiar property. When two pebbles meet in a single box, they annihilate and the energy released in the process creates a higher energy pebble in the box above. Of course, if there is already a pebble

---

[1]An alert reader will perhaps notice that what we call a *jump* in our every day life is also a gliding transition that takes a jumper, e.g., a cat, through a continuum of unstable configurations that may end eventually with the cat sitting stably on top of a table.

there, the newly created pebble and the previously inserted pebble annihilate too and an even higher energy pebble is created in the next box up.

Let us observe what is going to happen if we keep adding pebbles to the box at the bottom of the stack. When we place the first pebble there, the system looks as follows:
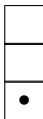
Now we add another pebble to the box at the bottom. The two pebbles annihilate and a new higher energy pebble is created in the middle box:
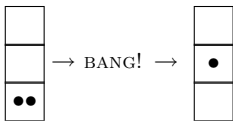
When we add a pebble again to the box at the bottom, nothing much happens, because there is no other pebble in it, and so the state of the register becomes:

But fireworks fly again, when we add a pebble to the box at the bottom this time:
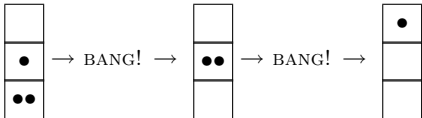
The first bang occurs because there are two pebbles in the box at the bottom of the stack. The pebbles annihilate and the energy released creates a

higher energy pebble in the middle box. But there is already a pebble there, so the two annihilate in the second bang, and a pebble of even higher energy is now created in the top box.

In summary the register has transitioned through the following stable states:
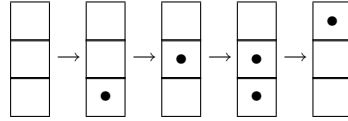
This register, as you see, is also a counter. Eventually we'll end up having pebbles in all three boxes. Adding a yet another pebble to the box at the bottom will release a chain reaction, which will clear all boxes and eject a very high energy pebble from the register altogether - this, therefore, is a modulo-7 counter.

creation and annihilation operators

The reason why this register is reminiscent of some quantum systems is because, first, its successive states are truly separate, without any in-betweens, i.e., pebbles don't move between boxes. Instead they simply vanish from a box, and if the energy released in the process is high enough, a new pebble emerges from nothingness in a higher energy box. This model bears some resemblance to quantum field theory, where particle states can be acted on by annihilation and creation operators. We will see similar formalism applied in the discussion of the Feynman quantum computer.

The second reason is that here we have a feature that resembles the Pauli exclusion principle, which states that no two fermions can coexist in the same state.

1.3   Fluctuating registers

The stable states of the register we have seen in the previous sections were all very well defined. For example the counter would go through the sequence of seven *stable* configurations:

$$\{000\} \rightarrow \{001\} \rightarrow \{010\} \rightarrow \{011\} \rightarrow \{100\} \rightarrow \{101\} \rightarrow \{110\}$$

Once a register would glide into one of these it would stay there, the values of its bits unchanging, until the next clock pulse would shift it to the next state. In general a 3-bit register can store numbers from 0 ({000}) through 7 ({111}) inclusive. Let us then focus on such a 3-bit register. It does not have to be a counter this time.

Suppose also that the register is afflicted by the following malady[2]. When set by some electronic procedure, let us say that there are some toggle switches to do this on the side of the package, to hold a binary number {101} its bits start to fluctuate randomly so that the register spends only 72% of the time in the {101} configuration, and 28% of the time in every other configuration, flickering at random between them. Suppose that the same happens when the register is set to hold other numbers, {000}, {001}, ..., {111} as well, i.e., the register ends up flickering between all possible configurations at random, but visits its *set* configuration 72% of the time. At first glance a register like this seems rather useless, but we could employ its fluctuations for example in Monte Carlo codes.

Now assume that after the register has been set, we cover its toggles with a masking tape and we are not allowed to ascertain or even describe the state of the register simply by looking at the toggles. Instead we have to resort to other means. The point of this exercise is not to torture the reader with needless equilibristics. Rather our aim is to prepare the reader for description of similar systems, which *do not* have any toggles at all.

The register exists in one of eight stable *states* associated with the setting of the toggles. Each state manifests by visiting a certain configuration more often than other configurations. However, this time we can no longer associate the *state* with the *configuration* as closely as we have done for the register that was not subject to random fluctuations. The state is now something more abstract, something that we can no longer associate with a simple single observation of the register. Instead we have to look at the register for a long time in order to identify its preferred configuration, and thus its *state*.

Let us introduce the following notation for these states of the fluctuating register:

$$\boldsymbol{p}_{\bar{0}} \quad \text{is the state that fluctuates around } \{000\}$$
$$\boldsymbol{p}_{\bar{1}} \quad \text{is the state that fluctuates around } \{001\}$$
$$\boldsymbol{p}_{\bar{2}} \quad \text{is the state that fluctuates around } \{010\}$$
$$\boldsymbol{p}_{\bar{3}} \quad \text{is the state that fluctuates around } \{011\}$$
$$\boldsymbol{p}_{\bar{4}} \quad \text{is the state that fluctuates around } \{100\}$$
$$\boldsymbol{p}_{\bar{5}} \quad \text{is the state that fluctuates around } \{101\}$$
$$\boldsymbol{p}_{\bar{6}} \quad \text{is the state that fluctuates around } \{110\}$$
$$\boldsymbol{p}_{\bar{7}} \quad \text{is the state that fluctuates around } \{111\}$$

---

[2]The *malady* may have been designed into the register on purpose.

Although we have named the states $\boldsymbol{p}_{\bar{0}}$ through $\boldsymbol{p}_{\bar{7}}$, we cannot at this early stage say what these objects really are. In order to endow the states of the fluctuating register with a mathematical structure we have to figure out how they can be measured and manipulated.

Observing a state of the fluctuating register

So, how can we ascertain which one of the eight states defined above the register is in, if we are not allowed to peek at the setting of its switches?

In order to do this we must observe the register for a long time writing down its observed configurations perhaps at random time intervals[3]. If the register is in the $\boldsymbol{p}_{\bar{5}}$ state, approximately 72% of the observations should return the {101} configuration with other results evenly spread over other configurations. Assuming we have made n_5 measurements of the register in total, $n^0{}_5$ observations would show the register in the {000} configuration, $n^1{}_5$ observations would show it in the {001} configuration, and so on for every other configuration, ending with $n^7{}_5$ for the {111} configuration[4]. We can now build the following column vector for which we would expect the following:

$$\begin{pmatrix} n^0{}_5/n_5 \approx 0.04 \\ n^1{}_5/n_5 \approx 0.04 \\ n^2{}_5/n_5 \approx 0.04 \\ n^3{}_5/n_5 \approx 0.04 \\ n^4{}_5/n_5 \approx 0.04 \\ n^5{}_5/n_5 \approx 0.72 \\ n^6{}_5/n_5 \approx 0.04 \\ n^7{}_5/n_5 \approx 0.04 \end{pmatrix}$$

We do not expect $n^6{}_5/n_5 = 0.04$ exactly, because, after all, the fluctuations of the register are random, but we do expect that we should get very close to 0.04 if n_5 is very large. For $n_5 \to \infty$ the ratios in the column vector above become *probabilities*. This lets us identify state $\boldsymbol{p}_{\bar{5}}$ with the column of probabilities of finding the register

The column of probabilities as the representation of a state of the fluctuating register

---

[3]An important assumption here is that we can observe the register without affecting its state.

[4]The superscripts 0 through 7 in $n^0{}_5$ through $n^7{}_5$ and also in $p^0{}_5$ through $p^7{}_5$ further down are *not* exponents. We do not raise n_5 (or p_5) to the powers of 0 through 7. They are just indexes, which say that, e.g., $n^4{}_5$ is the number of observations made on a register in state $\boldsymbol{p}_{\bar{5}}$ that found it in configuration {100} $\equiv 4$. There is a reason why we want this index to be placed in the superscript position rather than in the subscript position. This will be explained in more detail when we get to talk about forms and vectors in section 1.7 on page 28. If we ever need to exponentiate an object with a superscript index, e.g., $p^3{}_5$, we will enclose this object in brackets to distinguish between a raised index and an exponent, e.g., $\left(p^3{}_5\right)^2$.

in each of the eight possible configurations:

$$\boldsymbol{p}_{\bar{5}} \equiv \begin{pmatrix} p^0{}_5 = 0.04 \\ p^1{}_5 = 0.04 \\ p^2{}_5 = 0.04 \\ p^3{}_5 = 0.04 \\ p^4{}_5 = 0.04 \\ p^5{}_5 = 0.72 \\ p^6{}_5 = 0.04 \\ p^7{}_5 = 0.04 \end{pmatrix}$$

It is often convenient to think of the fluctuating register in terms of a *statistical ensemble*.

Suppose that instead of a single fluctuating register we have a very large number of static, non-fluctuating registers, of which 4% are in the {000} configuration, 4% are in the {001} configuration, 4% are in the {010} configuration, 4% are in the {011} configuration, 4% are in the {100} configuration, 72% are in the {101} configuration, 4% are in the {110} configuration and 4% are in the {111} configuration. Now let us put all the registers in a hat, mix them thoroughly, and draw at random n_5 registers from the hat. Of these $n^0{}_5$ will be in the {000} configuration, $n^1{}_5$ in the {001} configuration, ..., $n^6{}_5$ in the {110} configuration and $n^7{}_5$ in the {111} configuration. If the whole *ensemble* has been mixed well, we would expect that

$$
\begin{aligned}
n^0{}_5/n_5 &\approx 0.04 \\
n^1{}_5/n_5 &\approx 0.04 \\
n^2{}_5/n_5 &\approx 0.04 \\
n^3{}_5/n_5 &\approx 0.04 \\
n^4{}_5/n_5 &\approx 0.04 \\
n^5{}_5/n_5 &\approx 0.72 \\
n^6{}_5/n_5 &\approx 0.04 \\
n^7{}_5/n_5 &\approx 0.04
\end{aligned}
$$

Logically and arithmetically such an ensemble of static registers from which we *sample* n_5 registers is equivalent to a single randomly fluctuating register at which we *look* (without disturbing its overall condition) n_5 times.

The eight states our fluctuating register can be put in can be characterized by

A statistical ensemble of registers

Representing states of the register by columns of probabilities

the following column vectors of probabilities:

$$
\boldsymbol{p_{\bar{0}}} \equiv \begin{pmatrix} p^0{}_0 = 0.72 \\ p^1{}_0 = 0.04 \\ p^2{}_0 = 0.04 \\ p^3{}_0 = 0.04 \\ p^4{}_0 = 0.04 \\ p^5{}_0 = 0.04 \\ p^6{}_0 = 0.04 \\ p^7{}_0 = 0.04 \end{pmatrix}
\qquad
\boldsymbol{p_{\bar{1}}} \equiv \begin{pmatrix} p^0{}_1 = 0.04 \\ p^1{}_1 = 0.72 \\ p^2{}_1 = 0.04 \\ p^3{}_1 = 0.04 \\ p^4{}_1 = 0.04 \\ p^5{}_1 = 0.04 \\ p^6{}_1 = 0.04 \\ p^7{}_1 = 0.04 \end{pmatrix}
\qquad
\boldsymbol{p_{\bar{2}}} \equiv \begin{pmatrix} p^0{}_2 = 0.04 \\ p^1{}_2 = 0.04 \\ p^2{}_2 = 0.72 \\ p^3{}_2 = 0.04 \\ p^4{}_2 = 0.04 \\ p^5{}_2 = 0.04 \\ p^6{}_2 = 0.04 \\ p^7{}_2 = 0.04 \end{pmatrix}
$$

$$
\boldsymbol{p_{\bar{3}}} \equiv \begin{pmatrix} p^0{}_3 = 0.04 \\ p^1{}_3 = 0.04 \\ p^2{}_3 = 0.04 \\ p^3{}_3 = 0.72 \\ p^4{}_3 = 0.04 \\ p^5{}_3 = 0.04 \\ p^6{}_3 = 0.04 \\ p^7{}_3 = 0.04 \end{pmatrix}
\qquad
\boldsymbol{p_{\bar{4}}} \equiv \begin{pmatrix} p^0{}_4 = 0.04 \\ p^1{}_4 = 0.04 \\ p^2{}_4 = 0.04 \\ p^3{}_4 = 0.04 \\ p^4{}_4 = 0.72 \\ p^5{}_4 = 0.04 \\ p^6{}_4 = 0.04 \\ p^7{}_4 = 0.04 \end{pmatrix}
\qquad
\boldsymbol{p_{\bar{5}}} \equiv \begin{pmatrix} p^0{}_5 = 0.04 \\ p^1{}_5 = 0.04 \\ p^2{}_5 = 0.04 \\ p^3{}_5 = 0.04 \\ p^4{}_5 = 0.04 \\ p^5{}_5 = 0.72 \\ p^6{}_5 = 0.04 \\ p^7{}_5 = 0.04 \end{pmatrix}
$$

$$
\boldsymbol{p_{\bar{6}}} \equiv \begin{pmatrix} p^0{}_6 = 0.04 \\ p^1{}_6 = 0.04 \\ p^2{}_6 = 0.04 \\ p^3{}_6 = 0.04 \\ p^4{}_6 = 0.04 \\ p^5{}_6 = 0.04 \\ p^6{}_6 = 0.72 \\ p^7{}_6 = 0.04 \end{pmatrix}
\qquad
\boldsymbol{p_{\bar{7}}} \equiv \begin{pmatrix} p^0{}_7 = 0.04 \\ p^1{}_7 = 0.04 \\ p^2{}_7 = 0.04 \\ p^3{}_7 = 0.04 \\ p^4{}_7 = 0.04 \\ p^5{}_7 = 0.04 \\ p^6{}_7 = 0.04 \\ p^7{}_7 = 0.72 \end{pmatrix}
$$

We will call the probabilities that populate the arrays *fiducial measurements* and we will call the arrays of probabilities *fiducial vectors*[5] [26].

For every state $\boldsymbol{p_{\bar{i}}}$, $i = 0, 1, \ldots 7$, listed above we have that $p^0{}_i + p^1{}_i + p^2{}_i + p^3{}_i + p^4{}_i + p^5{}_i + p^6{}_i + p^7{}_i = 1$. This means that the probability of finding the register in any one of the configurations from $\{000\}$ through $\{111\}$ is 1. States $\boldsymbol{p_{\bar{i}}}$ that have this property are said to be *normalized*.

Given the collection of normalized states $\boldsymbol{p_{\bar{i}}}$ we can construct statistical ensembles with other values for probabilities p^0 through p^7 by *mixing* states $\boldsymbol{p_{\bar{i}}}$ in various proportions.

---

[5]The word *fiducial* in physics means an object or a system that is used as a standard of reference or measurement. It derives from the Latin word *fiducia*, which means *confidence* or *reliance*. Here we *rely* on the probability measurements in order to characterize the state of the system.

1.4 Mixtures and pure states

Suppose we have a very large number, N, of fluctuating registers affected by the malady discussed in the previous section. Suppose that N_0 of these have been put in state $\boldsymbol{p}_{\bar{0}}$ and the remaining $N - N_0 = N_3$ have been put in state $\boldsymbol{p}_{\bar{3}}$.

Now let us place all N registers into a hat and mix them thoroughly. We can draw them from the hat at random and look at their configuration but only once per register drawn. What probabilities should we expect for any possible register configuration in this ensemble? *Mixing registers in various states*

The easiest way to answer this question is to *expand* states $\boldsymbol{p}_{\bar{0}}$ and $\boldsymbol{p}_{\bar{3}}$ into their corresponding statistical ensembles and say that we have N_0 ensembles that correspond to state $\boldsymbol{p}_{\bar{0}}$ and N_3 ensembles that correspond to state $\boldsymbol{p}_{\bar{3}}$. In each $\boldsymbol{p}_{\bar{0}}$ ensemble we have $n^0{}_0$ registers out of n_0 in the $\{000\}$ configuration and in each $\boldsymbol{p}_{\bar{3}}$ ensemble we have $n^0{}_3$ registers out of n_3 in the $\{000\}$ configuration. So the total number of registers in the $\{000\}$ configuration is *Making use of statistical ensembles*

$$N_0 n^0{}_0 + N_3 n^0{}_3$$

The total number of registers after this expansion of states into ensembles is

$$N_0 n_0 + N_3 n_3$$

Therefore the probability of drawing a register in the $\{000\}$ configuration is going to be

$$\frac{N_0 n^0{}_0 + N_3 n^0{}_3}{N_0 n_0 + N_3 n_3}$$

in the limit $N \to \infty$, $n_0 \to \infty$ and $n_3 \to \infty$.

We can assume here, without any loss of generality, that we have an identical number of registers in the ensembles for $\boldsymbol{p}_{\bar{0}}$ and $\boldsymbol{p}_{\bar{3}}$, i.e., that $n_0 = n_3 = n$. Then

$$\frac{N_0 n^0{}_0 + N_3 n^0{}_3}{N_0 n_0 + N_3 n_3}$$
$$= \frac{N_0 n^0{}_0 + N_3 n^0{}_3}{n\,(N_0 + N_3)}$$
$$= \frac{N_0 n^0{}_0 + N_3 n^0{}_3}{nN}$$
$$= \frac{N_0}{N} \frac{n^0{}_0}{n} + \frac{N_3}{N} \frac{n^0{}_3}{n}$$

In the limit $n \to \infty$ and $N \to \infty$ this becomes:

$$P_0 p^0{}_0 + P_3 p^0{}_3$$

where P_0 is the probability of drawing a register in state $\boldsymbol{p}_{\bar{0}}$, P_3 is the probability of drawing a register in state $\boldsymbol{p}_{\bar{3}}$, $p^0{}_0$ is the probability that a register in state $\boldsymbol{p}_{\bar{0}}$ is observed in configuration $\{000\}$ and $p^0{}_3$ is the probability that a register in state $\boldsymbol{p}_{\bar{3}}$ is observed in configuration $\{000\}$.

Now assume that $P_0 = 0.3$ and that $P_3 = 0.7$. At this level (of probabilities pertaining to the mixture) we have that $P_0 + P_3 = 1$. What are the probabilities for each configuration in the mixture?

$$
\begin{aligned}
p^0 = P_0 p^0{}_0 + P_3 p^0{}_3 &= 0.3 \cdot 0.72 + 0.7 \cdot 0.04 = 0.244 \\
p^1 = P_0 p^1{}_0 + P_3 p^1{}_3 &= 0.3 \cdot 0.04 + 0.7 \cdot 0.04 = 0.04 \\
p^2 = P_0 p^2{}_0 + P_3 p^2{}_3 &= 0.3 \cdot 0.04 + 0.7 \cdot 0.04 = 0.04 \\
p^3 = P_0 p^3{}_0 + P_3 p^3{}_3 &= 0.3 \cdot 0.04 + 0.7 \cdot 0.72 = 0.516 \\
p^4 = P_0 p^4{}_0 + P_3 p^4{}_3 &= 0.3 \cdot 0.04 + 0.7 \cdot 0.04 = 0.04 \\
p^5 = P_0 p^5{}_0 + P_3 p^5{}_3 &= 0.3 \cdot 0.04 + 0.7 \cdot 0.04 = 0.04 \\
p^6 = P_0 p^6{}_0 + P_3 p^6{}_3 &= 0.3 \cdot 0.04 + 0.7 \cdot 0.04 = 0.04 \\
p^7 = P_0 p^7{}_0 + P_3 p^7{}_3 &= 0.3 \cdot 0.04 + 0.7 \cdot 0.04 = 0.04
\end{aligned}
$$

Observe that all probabilities p^i, $i = 1, \ldots, 7$, still add to one.

Fiducial vector and state for a mixture Using symbol $\boldsymbol{p}$ for the array of probabilities p^0 through p^7 we can write the above as follows:

$$
\boldsymbol{p} = P_0 \boldsymbol{p}_{\bar{0}} + P_3 \boldsymbol{p}_{\bar{3}}
$$

In general, assuming that we use all possible states in the mixture, we would have:

$$
\boldsymbol{p} = \sum_i P_i \boldsymbol{p}_{\bar{i}} \tag{1.1}
$$

Convexity i.e., the mixture state is a *linear combination* of its constituents. This linearity is restricted by two conditions, namely that $\sum_i P_i = 1$ and $\forall_i\ 0 \le P_i \le 1$. Linearity so restricted is called *convexity*, but we are going to show in section 1.6 that it can be extended to full linearity as long as what is on the left hand side of equation (1.1) is still a physically meaningful state.

It is easy to see that $\sum_k p^k = 1$ for the mixture:

$$
\begin{aligned}
\sum_k p^k &= \sum_k \sum_i P_i p^k{}_i = \sum_i P_i \sum_k p^k{}_i = \\
&= \sum_i P_i \cdot 1 = \sum_i P_i = 1
\end{aligned}
$$

States of fluctuating registers are mixtures of non-fluctuating states

States $\boldsymbol{p}_{\bar{0}}$ through $\boldsymbol{p}_{\bar{7}}$ are mixtures too. For example, state $\boldsymbol{p}_{\bar{3}}$, which is specified by the array:

$$\boldsymbol{p}_{\bar{3}} \equiv \begin{pmatrix} p^0{}_3 = 0.04 \\ p^1{}_3 = 0.04 \\ p^2{}_3 = 0.04 \\ p^3{}_3 = 0.72 \\ p^4{}_3 = 0.04 \\ p^5{}_3 = 0.04 \\ p^6{}_3 = 0.04 \\ p^7{}_3 = 0.04 \end{pmatrix}$$

can be thought of as a mixture of, say, 1,000,000 non-fluctuating registers of which 720,000 are in the {011} configuration at all times, and the remaining 280,000 registers are evenly spread over the remaining configurations, with 40,000 registers in each.

A register that is in a non-fluctuating {000} configuration can be still described in terms of a column vector of probabilities as follows:

Probabilistic description of non-fluctuating registers

$$\begin{pmatrix} p^0 = 1 \\ p^1 = 0 \\ p^2 = 0 \\ p^3 = 0 \\ p^4 = 0 \\ p^5 = 0 \\ p^6 = 0 \\ p^7 = 0 \end{pmatrix}$$

which states that the probability of finding this register in configuration {000} is 1, or, in other words, that the register spends 100% of its time in this configuration. We can use similar array representations for registers in non-fluctuating configurations {001} through {111}. These states, however, cannot be constructed by *mixing* other states, because probabilities cannot be negative, so there is no way that the zeros can be generated in linear combinations of non-zero coefficients, all of which represent some probabilities. States that are not mixtures are called *pure*. Only for pure states $\mid \boldsymbol{p} \rangle$ do we have that

Pure states

$$\sum_i p^i = 1 \quad \text{and} \quad \sum_i \left(p^i \right)^2 = 1$$

whereas for mixtures

$$\sum_i \left(p_i \right)^2 < 1$$

It is easy to see why this should be so. For $0 < p^i < 1$ we have that $0 < \left(p^i\right)^2 < p^i$. From this it follows that for such p^is $0 < \sum_i \left(p^i\right)^2 < \sum_i p^i = 1$. The equality $\sum_i \left(p^i\right)^2 = 1$ can therefore happen only if at least one $p^i = 1$, but since $\sum_i p^i = 1$ and none of the p_is can be negative, in this case all other p^is must be zero.

Fiducial representation of pure states

Let us introduce the following notation for the pure states:

$$
\boldsymbol{e}_0 \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad
\boldsymbol{e}_1 \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad
\boldsymbol{e}_2 \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
$$

$$
\boldsymbol{e}_3 \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad
\boldsymbol{e}_4 \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad
\boldsymbol{e}_5 \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}
$$

$$
\boldsymbol{e}_6 \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad
\boldsymbol{e}_7 \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}
$$

This time we have dropped the bars above the digits to emphasize that these *pure* states do not fluctuate and therefore the digits that represent them are exact and do not represent averages or most often encountered configurations.

The basis of pure states

Consider a randomly fluctuating register state $|\boldsymbol{p}\rangle$ specified by a vector of probabilities $\boldsymbol{p}$. Using straightforward array arithmetic we can express the array $\boldsymbol{p}$ in terms of $\boldsymbol{e}_i$ as follows

$$
\boldsymbol{p} = \sum_i p^i \boldsymbol{e}_i
$$

This expression is reminiscent of equation (1.1) on page 14, but here we have replaced fluctuating states on the right hand side with non-fluctuating *pure* states. And so, we have arrived at the following conclusion

> Every randomly fluctuating classical register state is a mixture of pure states.

The ability to decompose any randomly fluctuating register state $\boldsymbol{p}$ into a linear (to be exact, a convex) combination of other fluctuating or pure states suggests that we can think of the fluctuating states as belonging to a vector space in which the natural choice for the basis are the pure states[6]. But fluctuating states do not fill the space entirely, because only vectors for which

States of randomly fluctuating registers belong to a vector space

$$\sum_i p^i = 1 \quad \text{and} \quad \forall_i\, 0 \le p^i \le 1$$

are physical. Let us call the set of physically meaningful vectors in this space S.

Figure 1.4 shows set S for a 3-dimensional vector space that corresponds to a 2-bit modulo-3 fluctuating register, i.e., a register, for which the {11} configuration is unstable and flips the register back to {00}.

A set of physically meaningful states
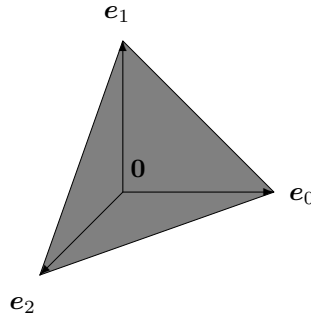
Figure 1.4: The set S in the 3-dimensional vector space that corresponds to a 2-bit modulo-3 fluctuating register is the gray triangle spanned by the ends of the three basis vectors of the space.

---

[6]We are going to *load* the term *basis* with an additional meaning soon, but what we have just called *the basis states* will remain such in the classical physics context even with this additional loading put upon them.

In this case S is the triangle spanned by the tips of the pure states

$$\boldsymbol{e}_0 \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\boldsymbol{e}_1 \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\boldsymbol{e}_2 \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

The null state   It is possible to fill the space between the triangle in figure 1.4 and the zero of the vector space by admitting the *null* state as a possible participant in the mixtures. A 3-bit register null state, for which we are going to use symbol $\boldsymbol{0}$, corresponds to the array of probabilities

$$\boldsymbol{0} \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The meaning of the null state is that a register in this state does not return any reading at all: we can say that it's broken, all its LEDs are off.

Consider the following mixture: 30% of the 3-bit registers are in the null state $\boldsymbol{0}$, 40% of the 3-bit registers are in the $\boldsymbol{e}_2$ state and the remaining 30% are in the $\boldsymbol{e}_4$ state. What are the coefficients p^i of the mixture? Let us expand the mixture into its statistical ensemble assuming for simplicity that the total number of registers in the ensemble is 100. 40 registers will then be in the {010} configuration and 30 registers will be in the {100} configuration. The remaining 30 registers will be in no readable configuration at all. The probability of drawing a register in the {010} configuration from the ensemble is therefore $p^2 = 40/100 = 0.4$. The probability of drawing a register in the {100} configuration from the ensemble is $p^4 = 30/100 = 0.3$. The probability

of drawing a register in any other configuration is 0. The resulting state of
the register is therefore

$$\boldsymbol{p} \equiv \begin{pmatrix} p^0 = 0.0 \\ p^1 = 0.0 \\ p^2 = 0.4 \\ p^3 = 0.0 \\ p^4 = 0.3 \\ p^5 = 0.0 \\ p^6 = 0.0 \\ p^7 = 0.0 \end{pmatrix}$$

The sum of all p^is is now equal 0.7, which is less than 1.

A state vector for which $\sum_i p^i < 1$ is said to be *unnormalized* as opposed
to a state vector for which $\sum_i p^i = 1$, which, as we have already remarked,
is said to be *normalized*. The addition of the null state **0** to the mixture
has the effect of "diluting" it. Because every unnormalized state vector must
have some admixture of the null state, it is clear that all unnormalized states
are mixtures. And conversely, states that are not mixtures and that are not
null either, i.e., pure states must be normalized. And then we also have the
body of states that are not pure, but are not diluted either: these states are
normalized and they are mixtures too.

Normalized and unnormalized states

Diluted mixtures

If null states are allowed then the corresponding set S is no longer restricted to
the surface of the triangle shown in figure 1.4. Instead the states fill the whole
volume of the tetrahedron between the triangle and the zero of the vector space,
including both the zero and the triangle.

The pure states and the null state are the *extremal* points of S. This observation
lets us arrive at the following definition of *pure* states:

Pure states as extremal points of S

Pure states correspond to extremal points of $S - \{\mathbf{0}\}$.

This definition will come handy in more complex situations and in richer theories,
in which we may not be able to draw a simple picture, or recognize that a given
state is pure by merely looking at its corresponding column of probabilities.

1.5   Basis states

In section 1.4 we stated that pure states were the natural choice for the basis of
the vector space in which physical states of fluctuating registers filled set S. This

assertion was based on the elementary algebra, which we used to decompose an arbitrary mixture into a linear combination of pure states:

$$\boldsymbol{p} = \sum_{i=0}^{7} p^i \boldsymbol{e}_i$$

This decomposition, in turn, was based on the representation of pure states by arrays of zeros and ones – arrays, which in the world of elementary algebra are commonly associated with basis vectors.

The "momentary glance" definition of a basis state

Here we are going to give a quite special physical meaning to what we are going to call the *basis state* throughout the remainder of this book and to what is also called the basis state in quantum mechanics.

> A *basis state* is the configuration of a randomly fluctuating register that can be ascertained by glancing at it momentarily.

Physical and canonical basis states

If we have a fluctuating register and *glance at it momentarily* we are not going to see it fluctuate. Instead, we are going to see this register in a quite specific frozen (albeit momentarily only) configuration, for example, {010}. If we glance at this register again a while later, the register may be in the {101} configuration[7]. We are going to call the non-fluctuating states that correspond to these configurations, i.e., $\boldsymbol{e}_2$ and $\boldsymbol{e}_5$, the (physical) basis states. Altogether we are going to have eight such basis states for the classical 3-bit randomly fluctuating register, assuming that all bits are allowed to fluctuate freely. The basis states, as defined here, are $\boldsymbol{e}_0$ through $\boldsymbol{e}_7$, which is what we have also called pure states, and what is also a *canonical* basis in the fiducial vector space.

We are going to use a special notation for such *momentarily glanced* basis vectors to distinguish between them and the canonical basis. Borrowing from the traditions of quantum mechanics, we'll denote them by $\mid \boldsymbol{e}_0 \rangle$ through $\mid \boldsymbol{e}_7 \rangle$.

Dimensionality and degrees of freedom

The number of vectors in the physical basis of the classical randomly fluctuating register, this number is also called the *dimensionality* of the system, is the same as the number of probabilities that are needed to describe the state. The number of probabilities is also called the number of degrees of freedom of the system. This follows clearly from how the probabilities have been defined: they are probabilities of finding the register in one of its specific configurations, which here we have identified with the basis states, because it is only these configurations (and not the fluctuating states) that we can actually see when we give the register a brief glance.

---

[7]Here again we make the assumption that glancing at the register does not affect its state.

Denoting the dimensionality of the system by N and the numbers of degrees of freedom by K, we can state that for the classical randomly fluctuating register

$$K = N \qquad (1.2)$$

Why should we distinguish at all between canonical and physical basis, especially since they appear to be exactly the same for the case considered so far? The answer is that they will not be the same in the quantum register case. What more, we will find there that

$$K = N^2 \qquad (1.3)$$

An astute reader may be tempted to ask the following question: what if I happen to catch the register in one of its *unstable* states it goes through when it *glides* between the configurations described by symbols {000} through {111}? Such a configuration does not correspond to any of the "basis vectors" we have defined in this section.

What to do with transitional configurations

This is indeed the case. Our probabilistic description does not cover the configurations of the continuum through which the register glides between its *basis states* at all. Instead we have focused entirely on the stable discrete configurations.

One of the ways to deal with the problem is to define more precisely what we mean by "momentarily". Assuming that the transition between the stable configurations of the register, the gliding phase, lasts δt we want "momentarily" to be much longer than δt. But "momentarily" must not be too long, because then we'll see the register switch during observation. If the register stays in any given configuration for Δt on average, then we want "momentarily" to be much shorter than Δt. And so we arrive at a somewhat more precise definition of "momentarily":

What do we mean by "momentarily"

$$\delta t \ll \text{"momentarily"} \ll \Delta t$$

We will see that this problem is not limited to classical registers. There is a similar condition imposed on quantum observations, although the dynamics of quantum observations is quite different. But it still takes a certain amount of time and effort to force a quantum system into its basis state. If the act of observation is too lightweight and too fast the quantum system will not "collapse" to the basis state and the measurement will be incomplete.

Another quite different way to deal with the problem is to include the continuum of configurations the register glides through between its stable states into the model, and allow register configurations such as {0.76 0.34 0.18}. We would then have to add probabilities (or probability densities) of finding the register in such a state to our measurements and our theory. The number of dimensions of the system would then skyrocket, but this does not necessarily imply that the system would become intractable.

How to include transitional states in the model

This solution also has its equivalent in the world of quantum physics. Detailed investigations of the spectrum of hydrogen atom revealed that its spectral fringes were split into very fine structure, and that additional splitting occurred in presence of electric and magnetic fields. To account for every observed feature of the spectrum physicists had to significantly enlarge the initially simple theory of hydrogen atom so as to incorporate various quantum electrodynamic corrections.

1.6 Functions and measurements on mixtures

How can we define a function on a randomly fluctuating register *state*?

There are various ways to do this. For example, we could associate certain values, $f_i \in \mathbb{R}$, with specific configurations of the register, i.e., $\{000\}, \ldots, \{111\}$, and then we could associate average of f_i over the statistical ensemble that corresponds to $\boldsymbol{p}$ with $f(\boldsymbol{p})$. This is a very physical way of doing things, since all that we can see as we glance at the fluctuating register every now and then are its various basis states, i.e., momentarily frozen configurations. If every one of these configurations is associated with some value f_i, what we're going to perceive in terms of f_i over a longer time, as the register keeps fluctuating, is an average value of f_i.

Functions on states as averages over the ensembles

Arbitrary functions on states

Another way to define a function f on a fluctuating register state would be to construct an arbitrary mapping of the form

$$S \ni \boldsymbol{p} \mapsto f\left(p^0, p^1, p^2, p^3, p^4, p^5, p^6, p^7\right) \in \mathbb{R}$$

where $\mathbb{R}$ stands for real numbers.

This second way is very general and it could be used to define quite complicated non-linear functions on coefficients p_i, for example:

$$S \ni \boldsymbol{p} \mapsto \left(p^0\right)^2 + p^1 p^3 - \sin\left(p^2 p^3 p^4\right) + e^{p^5 p^6 p^7} \in \mathbb{R}$$

A function like this could be implemented by an electronic procedure. For example, the procedure could observe the register for some time collecting statistics and building a fiducial vector for it. Once the vector is sufficiently well defined, the above operation would be performed on its content and the value of f delivered on output. Knowing value of f on the basis states would not in general help us evaluate f on an arbitrary state $\boldsymbol{p}$. Similarly, knowing values of f on the components of a mixture would not in general help us evaluate f on the mixture itself. In every case we would have to carry out full fiducial measurements for the whole mixture and then only evaluate the function.

On the other hand, the strategy outlined at the beginning of the section leads to functions which have very nice and simple properties and which also cover a very important special case: that of the fiducial vector of the mixture itself.

Consider a register that has a tiny tunable laser linked to its circuitry and the coupling between the laser and the configuration of the register is such that when the register is in configuration {000} the laser emits red light, when the register is in configuration {111} the laser emits blue light and when the register is in any of the intermediate configurations the laser emits light of some color that is between red and blue. Let the frequency of light emitted by the laser when the register is in state e_i (which corresponds directly to a specific configuration) be f_i. Let us then define the frequency function f on the basis states[8] as follows

$$f(e_i) \doteq f_i$$

We are going to extend this definition to an arbitrary mixture

$$p = \sum_i p^i e_i$$

by calculating the average value of f_i over the ensemble that corresponds to state p. Let us call this average value $\bar{f}$ and let us use the arithmetic mean formula to calculate it. If the ensemble comprises n registers then np^0 registers are in state e_0, np^1 registers are in state e_1, ..., and np^7 registers are in state e_7. The arithmetic mean of f_i over the ensemble is: *Arithmetic mean average over the ensemble*

$$
\begin{aligned}
\bar{f} &= \frac{1}{n}\left(np^0 f_0 + np^1 f_1 + \cdots + np^7 f_7\right) \\
&= \frac{n}{n}\sum_{i=0}^{7} p^i f_i = \sum_i p^i f_i \\
&= \sum_i p^i f(e_i)
\end{aligned}
$$

We can define that $f(p) \doteq \bar{f}$, which yields the following formula:

$$f(p) = f\left(\sum_i p^i e_i\right) = \sum_i p^i f(e_i)$$

---

[8]These are, in fact, physical basis states, because we want to associate the definition with register configurations that can be observed by glancing at the register momentarily. So we should really write here $\mid e_i \rangle$. But recall that for the classical register they are really the same, so we'll avoid a great deal of notational complexity by using the canonical basis instead.

Now let us take two arbitrary states[9]

$$\boldsymbol{p}_1 = \sum_i p^i{}_1 \boldsymbol{e}_i \quad \text{and} \quad \boldsymbol{p}_2 = \sum_i p^i{}_2 \boldsymbol{e}_i$$

and evaluate f on a mixture of these:

$$
\begin{aligned}
f\left(P_1\boldsymbol{p}_1 + P_2\boldsymbol{p}_2\right) &= f\left(P_1 \sum_i p^i{}_1 \boldsymbol{e}_i + P_2 \sum_i p^i{}_2 \boldsymbol{e}_i\rangle\right) \\
&= f\left(\sum_i \left(P_1 p^i{}_1 + P_2 p^i{}_2\right)\boldsymbol{e}_i\right) = \sum_i \left(P_1 p^i{}_1 + P_2 p^i{}_2\right) f\left(\boldsymbol{e}_i\right) \\
&= P_1 \sum_i p^i{}_1 f\left(\boldsymbol{e}_i\right) + P_2 \sum_i p^i{}_2 f\left(\boldsymbol{e}_i\right) \\
&= P_1 f\left(\sum_i p^i{}_1 \boldsymbol{e}_i\right) + P_2 f\left(\sum_i p^i{}_2 \boldsymbol{e}_i\right) \\
&= P_1 f\left(\boldsymbol{p}_1\right) + P_2 f\left(\boldsymbol{p}_2\right)
\end{aligned}
$$

In summary

$$f\left(P_1\boldsymbol{p}_1 + P_2\boldsymbol{p}_2\right) = P_1 f\left(\boldsymbol{p}_1\right) + P_2 f\left(\boldsymbol{p}_2\right) \tag{1.4}$$

Convexity of arithmetic mean on mixtures i.e., we find that f is *convex* on mixtures, or, in other words, it is linear on expressions of the form $P_1\boldsymbol{p}_1 + P_2\boldsymbol{p}_2$ where $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ belong to S, $0 \leq P_1 \leq 1$ and $0 \leq P_2 \leq 1$ and $P_1 + P_2 = 1$.

We can extend the definition of f to diluted mixtures by adding the following condition:

$$f\left(\boldsymbol{0}\right) = 0$$

Returning to our model where f_i is a frequency of light emitted by a tunable laser linked to a register configuration that corresponds to state $\boldsymbol{e}_i$, $f\left(\boldsymbol{p}\right)$ is the average frequency of light emitted by the laser as the register in state $\boldsymbol{p}$ fluctuates randomly through various configurations. Note that this average frequency may be different from all frequencies f_i that are actually observed. Frequency of light emitted by a broken register, $f\left(\boldsymbol{0}\right)$, is zero, i.e., in this case the laser does not emit anything.

Probability as a convex function on mixtures As we have already remarked, a very important class of functions that belong in this category are the probabilities associated with the mixture itself. Recall equation (1.1) on page 14, which we are going to rewrite here as follows:

$$\boldsymbol{p} = P_1\boldsymbol{p}_1 + P_2\boldsymbol{p}_2$$

---

[9]These are no longer the same states as our previously defined states $\boldsymbol{p}_{\bar{1}}$ and $\boldsymbol{p}_{\bar{2}}$: the bars above the subscripts are absent. They are simply two general mixture states.

Define a function on $\boldsymbol{p}$ that returns the i-th probability p^i. Let us call this function ω^i. Using this function and the above equation for the probability of a mixture we can write the following expression for p^i:

$$p^i = \omega^i\left(\boldsymbol{p}\right) = P_1\omega^i\left(\boldsymbol{p}_1\right) + P_2\omega^i\left(\boldsymbol{p}_2\right) = P_1 p^i{}_1 + P_2 p^i{}_2$$

If states $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ happen to be the canonical basis states, and if we have all of them in the mixture, we get

$$p^i = \omega^i\left(\boldsymbol{p}\right) = \omega^i\left(\sum_k p^k \boldsymbol{e}_k\right) = \sum_k p^k \omega^i\left(\boldsymbol{e}_k\right)$$

For this to make sense we must have that

$$\omega^i\left(\boldsymbol{e}_k\right) = \delta^i{}_k$$

where $\delta^i{}_k = 0$ for $i \neq k$ and $\delta^i{}_k = 1$ for $i = k$. *Kronecker delta*

Interpreting ω^i as an average value of something over the statistical ensemble that corresponds to $\boldsymbol{p}$, we can say that $p^i = \omega^i\left(\boldsymbol{p}\right)$ is the average frequency with which the basis state $\boldsymbol{e}_i$ is observed as we keep an eye on the randomly fluctuating register, e.g., 30 times out of 100 (on average) for $p^i = 30\%$.

The linear combinations of various states discussed so far were always restricted by the conditions

$$\forall_i \, 0 \leq p^i \leq 1 \quad \text{and} \quad \sum_i p^i = 1$$

or in case of diluted states $0 \leq \sum_i p^i < 1$. When we observed the linearity of function f defined as the arithmetic mean over statistical ensembles corresponding to states $\boldsymbol{p}$, it was also restricted to coefficients p^i or P_i (for mixtures of mixtures) satisfying the same conditions. So this *partial* linearity, the *convexity*, is not a full linearity, which should work also for $p^i > 1$ and for $p^i < 0$ – unless we can demonstrate that the former implies the latter.

So here we are going to demonstrate just this[10], i.e., that a function f *Convexity* defined as above, which has the property that *implies linearity*

$$f\left(P_1\boldsymbol{p}_1 + P_2\boldsymbol{p}_2\right) = P_1 f\left(\boldsymbol{p}_1\right) + P_2 f\left(\boldsymbol{p}_2\right) \tag{1.5}$$

where

$$0 \leq P_1 \leq 1 \quad \text{and} \quad 0 \leq P_2 \leq 1 \quad \text{and} \quad P_1 + P_2 = 1$$

---

[10]This proof follows [26] with some minor alterations.

is fully linear on S, i.e.,

$$f\left(\sum_i a_i \boldsymbol{p}_i\right) = \sum_i a_i f\left(\boldsymbol{p}_i\right) \tag{1.6}$$

where

$$\boldsymbol{p}_i \in S \quad \text{and} \quad \sum_i a_i \boldsymbol{p}_i \in S \quad \text{and} \quad a_i \in \mathbb{R}$$

i.e., a_i can be greater than 1 and they can be negative too.

Since we allow for the presence of the null state $\mathbf{0}$, assume that $\boldsymbol{p}_2 = \mathbf{0}$. The probability array (the fiducial vector) for $\mathbf{0}$ comprises zeros only and $f(\mathbf{0}) = 0$, hence equation (1.5) implies that in this case

$$f\left(P_1 \boldsymbol{p}_1\right) = P_1 f\left(\boldsymbol{p}_1\right) \tag{1.7}$$

Now replace P_1 with $1/\nu$ and $P_1 \boldsymbol{p}_1$ with $\boldsymbol{p}$:

$$f\left(\boldsymbol{p}\right) = \frac{1}{\nu} f\left(\nu \boldsymbol{p}\right)$$

or

$$\nu f\left(\boldsymbol{p}\right) = f\left(\nu \boldsymbol{p}\right) \tag{1.8}$$

Observe that in this new equation state $\boldsymbol{p}$ is unnormalized (or diluted) and $1 < \nu$. Combining equations (1.7) and (1.8) yields

$$f\left(a\boldsymbol{p}\right) = a f\left(\boldsymbol{p}\right) \tag{1.9}$$

for $0 \le a$ (including $1 < a$) and as long as $a\boldsymbol{p} \in S$, since otherwise the expression lacks physical meaning. But we can extend this expression beyond S from a purely algebraic point of view and this will come handy below.

Now let us consider an *arbitrary* linear combination of states that still delivers a state in S:

$$\boldsymbol{p} = \sum_i a_i \boldsymbol{p}_i$$

where, as above, we no longer restrict a_i: they can be negative and/or greater than 1 too. Let us divide coefficients a_i into negative and positive ones. Let us call the list of indexes i that yield $a_i < 0$ A_- and the list of indexes i that yield $a_i > 0$ A_+. This lets us rewrite the equation above as follows:

$$\boldsymbol{p} + \sum_{i \in A_-} |a_i| \boldsymbol{p}_i = \sum_{i \in A_+} a_i \boldsymbol{p}_i \tag{1.10}$$

Define

$$\nu = 1 + \sum_{i \in A_-} |a_i|$$

and let us divide both sides of equation (1.10) by ν:

$$\frac{1}{\nu}\boldsymbol{p} + \sum_{i \in A_-} \frac{|a_i|}{\nu}\boldsymbol{p}_i = \sum_{i \in A_+} \frac{a_i}{\nu}\boldsymbol{p}_i \qquad (1.11)$$

Observe that all coefficients on the left hand side of the equation, i.e.,

$$\frac{1}{\nu}, \frac{|a_i|}{\nu} \quad \text{for} \quad i \in A_-$$

are positive and add up to 1 (because we have defined ν so that they would add up to 1). Now define

$$\mu = \sum_{i \in A_+} \frac{a_i}{\nu}$$

Using μ we can additionally rewrite equation (1.11) as follows:

$$\frac{1}{\nu} \mid \boldsymbol{p}\rangle + \sum_{i \in A_-} \frac{|a_i|}{\nu} \mid \boldsymbol{p}_i\rangle = \mu \sum_{i \in A_+} \frac{a_i}{\mu\nu} \mid \boldsymbol{p}_i\rangle$$

Observe that all coefficients on the right hand side of this equation, i.e.,

$$\frac{a_i}{\mu\nu} \quad \text{for} \quad i \in A_+$$

are positive and add up to 1 (because we have defined μ so that they would add up to 1). We can now apply function f to both sides of this equation. Let us begin with the left hand side. Here we have a regular mixture, therefore:

$$f\left(\frac{1}{\nu}\boldsymbol{p} + \sum_{i \in A_-} \frac{|a_i|}{\nu}\boldsymbol{p}_i\right) = \frac{1}{\nu}f\left(\boldsymbol{p}\right) + \sum_{i \in A_-} \frac{|a_i|}{\nu}f\left(\boldsymbol{p}_i\right)$$

On the right hand side we first make use of equation (1.9) on page 26:

$$f\left(\mu \sum_{i \in A_+} \frac{a_i}{\mu\nu}\boldsymbol{p}_i\right) = \mu f\left(\sum_{i \in A_+} \frac{a_i}{\mu\nu}\boldsymbol{p}_i\right) = \cdots$$

and now we simply make use of the fact that what f acts on is a regular mixture and so:

$$\cdots = \mu \sum_{i \in A_+} \frac{a_i}{\mu\nu}f\left(\boldsymbol{p}_i\right)$$

Combining both sides yields:

$$\frac{1}{\nu}f\left(\boldsymbol{p}\right) + \sum_{i \in A_-} \frac{|a_i|}{\nu}f\left(\boldsymbol{p}_i\right) = \mu \sum_{i \in A_+} \frac{a_i}{\mu\nu}f\left(\boldsymbol{p}_i\right)$$

Let us finally multiply both sides of the equation by ν and let us cancel μ/μ on the right hand side to get:

$$f(\boldsymbol{p}) + \sum_{i \in A_-} |a_i| f(\boldsymbol{p}_i) = \sum_{i \in A_+} a_i f(\boldsymbol{p}_i)$$

which is the same as

$$f(\boldsymbol{p}) = \sum_i a_i f(\boldsymbol{p}_i)$$

in other words, f is, this time, fully linear on S.

Beyond   The arithmetic mean is not the only way in which function f can be extended
arithmetic mean from its definition on the basis states $\boldsymbol{e}_i$ to an arbitrary mixture $\boldsymbol{p}$. Instead of using arithmetic mean, $\langle \boldsymbol{f}, \boldsymbol{p} \rangle$, we could also use the generalized mean, which is defined by the formula:

$$\bar{f}_t = \left(\sum_i p^i \, (f_i)^t \right)^{1/t}$$

The generalized mean $\bar{f}_t$ becomes arithmetic mean for $t = 1$. It also becomes harmonic mean for $t = -1$ and geometric mean for $t \to 0$. For very large values of t $\bar{f}_t \approx \max_i f_i$ and for very large negative values of t $\bar{f}_t \approx \min_i f_i$.

There may indeed be situations in electronics and physics when, e.g., harmonic mean or geometric mean are more appropriate ways of extending f to the mixture. But f so defined would not be linear and therefore would not mix in direct proportion to the abundance of various components in the mixture.

1.7   Forms and Vectors

Linear functions on a vector space are called *forms*. Such functions can be thought of as mirror images of vectors they operate on[11] and they form a vector space of their own too.

Whenever a *form* $\boldsymbol{f}$, which represents a measurement on a mixture, encounters a vector $\boldsymbol{p}$, which represents a mixture, they get together in an explosive union $\langle \boldsymbol{f}, \boldsymbol{p} \rangle$ which delivers a number $f(\boldsymbol{p})$:

$$\langle \boldsymbol{f}, \boldsymbol{p} \rangle \doteq f(\boldsymbol{p})$$

---

[11]A great and quite painless introduction to vectors, forms and tensors can be found in [39], Part II, "Physics in Flat Spacetime", sections 2 and 3.

Using this notation we can restate the linearity of function f, equation (1.6), as follows:

$$\langle \boldsymbol{f}, \sum_i a_i \boldsymbol{p}_i \rangle = \sum_i a_i \langle \boldsymbol{f}, \boldsymbol{p}_i \rangle$$

In practical computations we often identify a state vector $\boldsymbol{p}$ with a column of numbers (probabilities), although much can be said and even proven about state vectors without using this particular representation. Similarly, a form $\boldsymbol{f}$ can be identified with a *row* of numbers. For example, a form that returns the first component of a vector would have the following row-of-numbers representation:

$$\boldsymbol{\omega}^0 \equiv (1, 0, 0, 0, 0, 0, 0, 0)$$

When a form like this is attached to a vector $\boldsymbol{p}$ the following results:

$$\langle \boldsymbol{\omega}^0, \boldsymbol{p} \rangle = (1, 0, 0, 0, 0, 0, 0, 0) \cdot \begin{pmatrix} p^0 \\ p^1 \\ p^2 \\ p^3 \\ p^4 \\ p^5 \\ p^6 \\ p^7 \end{pmatrix} = p^0$$

where the dot between the row and the column stands for matrix multiplication. This form represents a measurement of probability that the system is in configuration $\{000\}$.

The following listing introduces a canonical basis in the space of forms:

$$
\begin{aligned}
\boldsymbol{\omega}^0 &\equiv (1, 0, 0, 0, 0, 0, 0, 0) \\
\boldsymbol{\omega}^1 &\equiv (0, 1, 0, 0, 0, 0, 0, 0) \\
\boldsymbol{\omega}^2 &\equiv (0, 0, 1, 0, 0, 0, 0, 0) \\
\boldsymbol{\omega}^3 &\equiv (0, 0, 0, 1, 0, 0, 0, 0) \\
\boldsymbol{\omega}^4 &\equiv (0, 0, 0, 0, 1, 0, 0, 0) \\
\boldsymbol{\omega}^5 &\equiv (0, 0, 0, 0, 0, 1, 0, 0) \\
\boldsymbol{\omega}^6 &\equiv (0, 0, 0, 0, 0, 0, 1, 0) \\
\boldsymbol{\omega}^7 &\equiv (0, 0, 0, 0, 0, 0, 0, 1)
\end{aligned}
$$

As we did with columns of probabilities, we shall call the rows and symbols such *Fiducial forms*

as $\boldsymbol{\omega}^i$ *fiducial* forms or *fiducial measurements*, because they act on fiducial vectors.

The basis forms $\boldsymbol{\omega}^i$ satisfy the following relation:

$$\langle \boldsymbol{\omega}^i, \boldsymbol{e}_j \rangle = \delta^i{}_j$$

As we can express a mixture $\boldsymbol{p}$ in terms of the basis states

$$\boldsymbol{p} = \sum_i p^i \boldsymbol{e}_i$$

a form $\boldsymbol{f}$ can be expressed in terms of the basis forms $\boldsymbol{\omega}^i$ defined above as follows:

$$\boldsymbol{f} = \sum_i f_i \boldsymbol{\omega}^i$$

The action of $\boldsymbol{f}$ on $\boldsymbol{p}$ now becomes:

$$\begin{aligned}
\langle \boldsymbol{f}, \boldsymbol{p} \rangle &= \left\langle \sum_i f_i \boldsymbol{\omega}^i, \sum_j p^j \boldsymbol{e}_j \right\rangle \\
&= \sum_i \sum_j f_i p^j \langle \boldsymbol{\omega}^i, \boldsymbol{e}_j \rangle = \sum_i \sum_j f_i p^j \delta^i{}_j \\
&= \sum_i f_i p^i
\end{aligned}$$

The physical meaning of $\langle \boldsymbol{f}, \boldsymbol{p} \rangle$ is the arithmetic mean of f_i over the statistical ensemble represented by the mixture coefficients p^i, where f_i are the values f assumes on the basis states $\boldsymbol{e}_i$.

About subscripts
and superscripts
Observe the following typographic convention. Vectors and forms are typeset using bold font, whereas vector and form coefficients are typeset using light font. Whenever possible we will reserve small Latin letters, e.g., $\boldsymbol{v}$, $\boldsymbol{e}$, for vectors and small Greek letters, e.g., $\boldsymbol{\eta}$, $\boldsymbol{\omega}$, for forms – though sometimes, as we have just done with form $\boldsymbol{f}$, we will break this convention. Basis vectors, e.g., $\boldsymbol{e}_i$ are numbered with subscripts. Basis forms, e.g., $\boldsymbol{\omega}^i$ are numbered with superscripts. On the other hand, vector coefficients, e.g., p^i are numbered with superscripts and form coefficients, e.g., f_i are numbered with subscripts. This way whenever there is a summation in expressions such as:

$$\begin{aligned}
\boldsymbol{p} &= \sum_i p^i \boldsymbol{e}_i \\
\boldsymbol{f} &= \sum_i f_i \boldsymbol{\omega}^i \\
\langle \boldsymbol{f}, \boldsymbol{p} \rangle &= \sum_i f_i p^i
\end{aligned}$$

the summation runs over indexes of which one is always down and the other one is always up. Such summation is called *contraction*. One says, for example, about *contraction on index i*.

This convention is very useful, because it constantly reminds us about what various objects we work with actually are. It helps debug form and vector expressions too. For example, an expression such as $\sum_i f_i \boldsymbol{e}_i$ should attract our suspicions, because it suggests that we are trying to use form coefficients in order to construct a vector.

This typographic convention leads to the so called *summation convention*, which states that whenever there is an expression with two identical indexes, of which one is up and the other one down, summation should be assumed. For example

$$
\begin{aligned}
f_i p^i &\equiv \sum_i f_i p^i \\
p^i \boldsymbol{e}_i &\equiv \sum_i p^i \boldsymbol{e}_i \\
f_i \boldsymbol{\omega}^i &\equiv \sum_i f_i \boldsymbol{\omega}^i
\end{aligned}
$$

The summation convention is very handy in tensor calculus, where geometric and dynamic objects may be endowed with several subscripts and superscripts. We will not make much use of it though, and will always state so explicitly, when we do.

The placement of indexes on form and vector coefficients is not just a matter of esthetics, convenience and debugging. It reflects transformation properties of these objects too.

Suppose that instead of decomposing vector $\boldsymbol{v}$ in basis $\boldsymbol{e}_i$, $\boldsymbol{v} = \sum_i v^i \boldsymbol{e}_i$, we were to decompose it in another basis, say, $\boldsymbol{e}_{i'}$. Basis vectors $\boldsymbol{e}_{i'}$ are not the same as $\boldsymbol{e}_i$, the prime on the index i' matters, but they are all linearly independent as basis vectors should be. Suppose also that we find another basis in the form space, $\boldsymbol{\omega}^{i'}$ such that $\langle \boldsymbol{\omega}^{i'}, \boldsymbol{e}_{j'} \rangle = \delta^{i'}{}_{j'}$. Vector coefficients in the new basis $\boldsymbol{e}_{i'}$ can be found by using the basis in the form space, namely

$$
v^{i'} = \langle \boldsymbol{\omega}^{i'}, \boldsymbol{v} \rangle
$$

Since both $\boldsymbol{e}_i$ and $\boldsymbol{e}_{i'}$ are the bases of linearly independent vectors, there must be a linear transformation that converts one basis into another one. Let us call the coefficients of this transformation $\Lambda_{i'}{}^j$. The transformation rule for the basis vectors is then:

$$
\boldsymbol{e}_{i'} = \sum_j \Lambda_{i'}{}^j \boldsymbol{e}_j
$$

We should expect a similar transformation for the forms

$$
\boldsymbol{\omega}^{i'} = \sum_j \boldsymbol{\omega}^j \Lambda_j{}^{i'}
$$

We do not assume that $\Lambda_{i'}{}^{j}$ and $\Lambda_{j}{}^{i'}$ are the same: the typographic placement of primed and unprimed indexes warns us that they may be different. But they are related. This is easy to see by invoking the rules $\langle \boldsymbol{\omega}^i, \boldsymbol{e}_j \rangle = \delta^i{}_j$ and $\langle \boldsymbol{\omega}^{i'}, \boldsymbol{e}_{j'} \rangle = \delta^{i'}{}_{j'}$:

$$\delta^{i'}{}_{j'} = \langle \boldsymbol{\omega}^{i'}, \boldsymbol{e}_{j'} \rangle = \left\langle \sum_k \boldsymbol{\omega}^k \Lambda_k{}^{i'}, \sum_l \Lambda_{j'}{}^l \boldsymbol{e}_l \right\rangle$$

$$= \sum_k \sum_l \Lambda_k{}^{i'} \Lambda_{j'}{}^l \langle \boldsymbol{\omega}^k, \boldsymbol{e}_l \rangle = \sum_k \sum_l \Lambda_k{}^{i'} \Lambda_{j'}{}^l \delta^k{}_l$$

$$= \sum_k \Lambda_k{}^{i'} \Lambda_{j'}{}^k$$

This tells us that matrices $\left\| \Lambda_k{}^{i'} \right\|$ and $\left\| \Lambda_{j'}{}^k \right\|$ are inverses of each other.

Now we can turn back to transformation properties of vector and form coefficients. We can easily see that for vectors we have the following:

$$v^{i'} = \left\langle \boldsymbol{\omega}^{i'}, \boldsymbol{v} \right\rangle = \left\langle \sum_j \boldsymbol{\omega}^j \Lambda_j{}^{i'}, \boldsymbol{v} \right\rangle$$

$$= \sum_j \Lambda_j{}^{i'} \left\langle \boldsymbol{\omega}^j, \boldsymbol{v} \right\rangle = \sum_j v^j \Lambda_j{}^{i'}$$

On the other hand, we get a different relation for forms:

$$\eta_{i'} = \langle \boldsymbol{\eta}, \boldsymbol{e}_{i'} \rangle = \left\langle \boldsymbol{\eta}, \sum_j \Lambda_{i'}{}^j \boldsymbol{e}_j \right\rangle$$

$$= \sum_j \Lambda_{i'}{}^j \langle \boldsymbol{\eta}, \boldsymbol{e}_j \rangle = \sum_j \Lambda_{i'}{}^j \eta_j$$

We see that form and vector coefficients transform in opposite directions. Vector coefficients (index is up) transform like form basis (their index is up too) and form coefficients (index down) transform like vector basis (their index is down too). This is actually very good, because this means that expressions such as $\sum_i \eta_i v^i$ don't transform at all. Transformations of η_i and v^i cancel each other, so that the resulting scalar $\langle \boldsymbol{\eta}, \boldsymbol{v} \rangle$ is *independent* of the choice of vector and form bases.

1.8 Transformations of mixtures

Suppose we have a hat full of randomly mixed 3-bit registers in various *static* configurations. Such a statistical ensemble is equivalent to a single randomly fluctuating

register in some *state* $\boldsymbol{p}$. If there is a total of N registers in the hat, the abundances of registers in specific configurations are $Np^0, Np^1, \ldots, Np^7$.

Let us now call a local Cinderella, there is always bound to be one in a nearby *Cinderella* opera house, and let us give her the following ungrateful task. She should draw *transformation* the registers out of the hat one by one. Whenever she draws a register in state $\boldsymbol{e}_0$ she should tweak its toggles so as to change its state to $\boldsymbol{p}_0$ and then place it in another hat. Whenever she draws a register in state $\boldsymbol{e}_1$ she should tweak it so as to change its state to $\boldsymbol{p}_1$ and then place it in the other hat too, and so on for the remaining states. In other words, she should perform the following transformation on the *whole* ensemble:

$$
\begin{aligned}
\boldsymbol{e}_0 &\rightarrow \boldsymbol{p}_0 \\
\boldsymbol{e}_1 &\rightarrow \boldsymbol{p}_1 \\
\boldsymbol{e}_2 &\rightarrow \boldsymbol{p}_2 \\
\boldsymbol{e}_3 &\rightarrow \boldsymbol{p}_3 \\
\boldsymbol{e}_4 &\rightarrow \boldsymbol{p}_4 \\
\boldsymbol{e}_5 &\rightarrow \boldsymbol{p}_5 \\
\boldsymbol{e}_6 &\rightarrow \boldsymbol{p}_6 \\
\boldsymbol{e}_7 &\rightarrow \boldsymbol{p}_7
\end{aligned}
$$

where $\boldsymbol{p}_i = \sum_j p^j{}_i \boldsymbol{e}_j$.

What Cinderella is going to end up with in the second hat, after the whole operation is finished, is another mixture. Let us now draw a register from the second hat, this time it is going to be a fluctuating one, and glance at it momentarily. What are the probabilities of seeing $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_7$?

To answer this question we shall expand, as we did on previous occasions, states $\boldsymbol{p}_0$ through $\boldsymbol{p}_7$ into their statistical ensembles, remembering that we are going to have

Np^0	ensembles that correspond to	$\boldsymbol{p}_0$
Np^1	ensembles that correspond to	$\boldsymbol{p}_1$
Np^2	ensembles that correspond to	$\boldsymbol{p}_2$
Np^3	ensembles that correspond to	$\boldsymbol{p}_3$
Np^4	ensembles that correspond to	$\boldsymbol{p}_4$
Np^5	ensembles that correspond to	$\boldsymbol{p}_5$
Np^6	ensembles that correspond to	$\boldsymbol{p}_6$

$$Np^7 \qquad \text{ensembles that correspond to} \qquad \boldsymbol{p}_7$$

Without loss of generality we can assume that each ensemble that corresponds to $\boldsymbol{p}_0$ through $\boldsymbol{p}_7$ comprises the same number of non-fluctuating registers. Let's call this number n. Consequently, in an ensemble that corresponds to $\boldsymbol{p}_0$ we are going to have

$$np^0{}_0 \qquad \text{registers in state} \qquad \boldsymbol{e}_0$$
$$np^1{}_0 \qquad \text{registers in state} \qquad \boldsymbol{e}_1$$
$$np^2{}_0 \qquad \text{registers in state} \qquad \boldsymbol{e}_2$$
$$np^3{}_0 \qquad \text{registers in state} \qquad \boldsymbol{e}_3$$
$$np^4{}_0 \qquad \text{registers in state} \qquad \boldsymbol{e}_4$$
$$np^5{}_0 \qquad \text{registers in state} \qquad \boldsymbol{e}_5$$
$$np^6{}_0 \qquad \text{registers in state} \qquad \boldsymbol{e}_6$$
$$np^7{}_0 \qquad \text{registers in state} \qquad \boldsymbol{e}_7$$

and similarly for the other ensembles. The total number of registers in state $\boldsymbol{e}_0$ in all the ensembles is going to be:

$$Np^0np^0{}_0 + Np^1np^0{}_1 + Np^2np^0{}_2 + Np^3np^0{}_3$$
$$+ Np^4np^0{}_4 + Np^5np^0{}_5 + Np^6np^0{}_6 + Np^7np^0{}_7$$
$$= Nn\sum_{i=0}^{7} p^i p^0{}_i$$

The total number of all registers in all the ensembles is going to be Nn, because we had N registers in the first hat and after tweaking the toggles each register got "expanded" into an ensemble of n registers. Let the state of the mixture in the second hat be called $\boldsymbol{q}$ with coefficients $q^0, q^1, \ldots, q^7$. Making use of the above formula we have for q^0:

$$q^0 = \frac{1}{Nn} Nn \sum_{i=0}^{7} p^i p^0{}_i = \sum_{i=0}^{7} p^0{}_i p^i$$

We have reversed here the order of p_i and p_{0i} for a purely cosmetic reason. Similarly we can write the following equations for the remaining coefficients:

$$q^1 = \sum_{i=0}^{7} p^1{}_i p^i \quad q^2 = \sum_{i=0}^{7} p^2{}_i p^i \quad q^3 = \sum_{i=0}^{7} p^3{}_i p^i$$

$$q^4 = \sum_{i=0}^{7} p^4{}_i p^i \quad q^5 = \sum_{i=0}^{7} p^5{}_i p^i \quad q^6 = \sum_{i=0}^{7} p^6{}_i p^i$$

$$q^7 = \sum_{i=0}^{7} p^7{}_i p^i$$

All of this can be written more concisely as follows:

$$q^j = \sum_{i} p^j{}_i p^i$$

We can also rewrite this formula using vector and form notation:

$$\boldsymbol{q} = \sum_{j} q^j \boldsymbol{e}_j = \sum_{j} \sum_{i} p^j{}_i \boldsymbol{e}_j p^i = \sum_{j} \sum_{i} p^j{}_i \boldsymbol{e}_j \langle \boldsymbol{\omega}^i, \boldsymbol{p} \rangle$$

$$= \left\langle \sum_{j} \sum_{i} p^j{}_i \boldsymbol{e}_j \otimes \boldsymbol{\omega}^i, \boldsymbol{p} \right\rangle = \langle \boldsymbol{P}, \boldsymbol{p} \rangle$$

where we have defined

$$\boldsymbol{P} \doteq \sum_{j} \sum_{i} p^j{}_i \boldsymbol{e}_j \otimes \boldsymbol{\omega}^i$$

The symbol $\otimes$ is called a *tensor product* and all that it means here is that we put   *Tensor product*
$\boldsymbol{e}_j$ and $\boldsymbol{\omega}^i$ together next to each other *typographically*. If form $\boldsymbol{\omega}^i$ finds a vector to
prey on, e.g., $\boldsymbol{p}$, it vanishes together with it leaving a number $\langle \boldsymbol{\omega}^i, \boldsymbol{p} \rangle$ behind and
all that is left on the paper then is $\boldsymbol{e}_j$ multiplied by that number.

Symbol $\boldsymbol{P}$ denotes an *operator* that describes the transformation of the ensemble
performed by Cinderella. Observe that every j-th term of this operator, $\sum_i p^j{}_i \boldsymbol{\omega}^i$
is a form. The operator can therefore be thought of as eight forms arranged so
that they together transform one state vector, $\boldsymbol{p}$, into another one, $\boldsymbol{q}$. Each of the
forms is a convex function, but, as we have already seen, it is also fully linear on
S. Consequently $\boldsymbol{P}$, the collection of the forms, is linear too.

We will usually adhere throughout this text to the typographic convention
which uses capital bold letters, like $\boldsymbol{P}$, for operators and other complex objects   *Typographic*
that have one or more tensor products inside them, although we will deviate   *conventions*
from it in some cases, where tradition dictates that, e.g., a metric tensor   *pertaining to*
should be denoted by $\boldsymbol{g}$.   *operators*

Another typographic convention drops brackets $\langle$ and $\rangle$ when describing
the action of an operator on a vector. And so, instead of writing $\langle \boldsymbol{P}, \boldsymbol{p} \rangle$ we can
write simply $\boldsymbol{P}\boldsymbol{p}$. This has the additional benefit of translating naturally into
a matrix (representing the $\boldsymbol{P}$) times a column (representing the $\boldsymbol{p}$) expression.

Convexity of
Cinderella
transformations

Although every Cinderella transformation like the one discussed here is going to be linear, not every linear transformation that we can apply to S is going to be a valid Cinderella transformation. Rotations and reflections, for example, are linear, but here they would rotate or reflect some of the states out of S. Cinderella transformations on the other hand keep everything within S because of the following three conditions, which, all together, imply convexity of $\boldsymbol{P}$:

$$\forall_{ji} 0 \leq p^j{}_i \leq 1$$

$$\forall_j \sum_i p^j{}_i = 1$$

$$\forall_i \sum_j p^j{}_i = 1$$

Reversibility of
Cinderella
transformations

Can Cinderella transformations be reversed? There is a very interesting answer to this question: the only reversible Cinderella transformations are permutations on pure states, i.e., if a reversible Cinderella transformation is applied to a pure state, another pure state must come out on output. Since reversibility also implies that no two different pure states may be converted to the same output state, the only possibility we are left with is a permutation of pure states.

The way to see that a reversible Cinderella transformation must convert a pure state into another pure state is as follows. Let the transformation in question be called $\boldsymbol{C}$. If it is possible for $\boldsymbol{C}$ to convert a pure state to a mixture then we would have that:

$$\boldsymbol{C}\boldsymbol{e}_i = \sum_k c^k \boldsymbol{e}_k$$

Since $\boldsymbol{C}$ is reversible, $\boldsymbol{C}^{-1}$ exists and we can apply it to both sides of the equation above:

$$\boldsymbol{C}^{-1}\boldsymbol{C}\boldsymbol{e}_i = \boldsymbol{e}_i = \sum_k c^k \boldsymbol{C}^{-1}\boldsymbol{e}_k$$

But this says that $\boldsymbol{e}_i$ is a mixture[12], which it is not. Hence we must conclude that $\boldsymbol{C}$ cannot convert $\boldsymbol{e}_k$ to a mixture.

Beyond linearity

We can always think of a quite general transformation on a mixture defined by the following formula:

$$\boldsymbol{q} = \sum_i q^i \left(p^0, p^1, \ldots, p^7 \right) \boldsymbol{e}_i$$

---

[12]Note that $\boldsymbol{C}^{-1}\boldsymbol{e}_k$ must be linearly independent states because otherwise $\boldsymbol{C}$ wouldn't be reversible.

where $q^i \left(p^0, p^1, \ldots, p^7\right)$ are some arbitrary, possibly nonlinear, real-valued functions that convert coefficients p^i, $i = 0, 1, \ldots, 7$ into real numbers between 0 and 1. The only additional condition we would impose on functions q^i would be that

$$\forall_{\boldsymbol{p} \in S} \sum_i q^i \left(\boldsymbol{p}\right) \leq 1$$

where the ≤ 1 condition would cover the option of generating diluted mixtures. How could we implement such a general transformation? It would not be sufficient to provide Cinderella with a prescription such as before – this, as we have seen, would result in a linear transformation. In order to generate $\boldsymbol{q}$ Cinderella would have to empty the first hat entirely, counting abundances for each configuration. Then only, having collected sufficient fiducial statistics to ascertain the state of the *whole* mixture $\boldsymbol{p}$, she could sit down and calculate q^i for $i = 0, 1, \ldots, 7$. Having done this she could then set switches on each register to generate $\boldsymbol{q}$ and only then she would place the register in the second hat.

Nonlinear transformations cannot be implemented without the full knowledge of the state

This transformation would be quite different physically from a Cinderella transformation. A Cinderella transformation can be implemented on the go: we don't have to have the complete knowledge of the mixture in order to begin processing the registers. The prescription allows us to perform the transformation on each register separately and things will still add up to $\boldsymbol{q} = \boldsymbol{Pp}$.

1.9 Composite systems

In this section we are going to have a closer look at what happens when we combine smaller randomly fluctuating registers into a larger one.

Consider two 2-bit randomly fluctuating registers. A 2-bit randomly fluctuating register is described by states that belong to the $2^2 = 4$ dimensional vector space. The canonical basis vectors in this space are $\boldsymbol{e}_0 \equiv \{00\}$, $\boldsymbol{e}_1 \equiv \{01\}$, $\boldsymbol{e}_2 \equiv \{10\}$ and $\boldsymbol{e}_3 \equiv \{11\}$. A vector space that contains the register obtained by combining the two 2-bit registers is $4 \times 4 = 16$ dimensional. Let us call the two 2-bit registers A and B and let us label the basis states that refer to these two registers with indexes A and B too. And so we have the following basis states for register A:

$$\{00\}_A, \{01\}_A, \{10\}_A, \{11\}_A$$

and similarly for register B:

$$\{00\}_B, \{01\}_B, \{10\}_B, \{11\}_B$$

What will the basis states look like of a system made by placing registers A and B next to each other, so that they form a 4-bit register? Since we are not going to do anything special to these two registers, other than just place them next to each other, whenever we give them a brief glance as they keep fluctuating, we're going to see one of the following:

Basis of the combined register system

$$\{00\}_A\{00\}_B \qquad \{00\}_A\{01\}_B$$
$$\{00\}_A\{10\}_B \qquad \{00\}_A\{11\}_B$$
$$\{01\}_A\{00\}_B \qquad \{01\}_A\{01\}_B$$
$$\{01\}_A\{10\}_B \qquad \{01\}_A\{11\}_B$$
$$\{10\}_A\{00\}_B \qquad \{10\}_A\{01\}_B$$
$$\{10\}_A\{10\}_B \qquad \{10\}_A\{11\}_B$$
$$\{11\}_A\{00\}_B \qquad \{11\}_A\{01\}_B$$
$$\{11\}_A\{10\}_B \qquad \{11\}_A\{11\}_B$$

And so these pairs must be then the basis state vectors of the combined system. We can replace binary digits in the curly brackets with our symbolic notation for basis vectors as follows

$$\{01\}_A\{10\}_B \rightarrow \boldsymbol{e}_{1A}\boldsymbol{e}_{2B}$$

where we have simply placed $\boldsymbol{e}_{1A}$ and $\boldsymbol{e}_{2B}$ next to each other on the sheet of paper. But we have already seen something very similar when we defined an operator in terms of the tensor product $\otimes$. So just to avoid a possible confusion and emphasize that we do not really multiply these vectors by each other but merely write then next to each other, let us use the same symbol here:

$$\{01\}_A\{10\}_B \rightarrow \boldsymbol{e}_{1A} \otimes \boldsymbol{e}_{2B}$$

The basis of the combined register system now becomes:

$$\boldsymbol{e}_{0A} \otimes \boldsymbol{e}_{0B} \qquad \boldsymbol{e}_{0A} \otimes \boldsymbol{e}_{1B}$$
$$\boldsymbol{e}_{0A} \otimes \boldsymbol{e}_{2B} \qquad \boldsymbol{e}_{0A} \otimes \boldsymbol{e}_{3B}$$
$$\boldsymbol{e}_{1A} \otimes \boldsymbol{e}_{0B} \qquad \boldsymbol{e}_{1A} \otimes \boldsymbol{e}_{1B}$$
$$\boldsymbol{e}_{1A} \otimes \boldsymbol{e}_{2B} \qquad \boldsymbol{e}_{1A} \otimes \boldsymbol{e}_{3B}$$
$$\boldsymbol{e}_{2A} \otimes \boldsymbol{e}_{0B} \qquad \boldsymbol{e}_{2A} \otimes \boldsymbol{e}_{1B}$$
$$\boldsymbol{e}_{2A} \otimes \boldsymbol{e}_{2B} \qquad \boldsymbol{e}_{2A} \otimes \boldsymbol{e}_{3B}$$
$$\boldsymbol{e}_{3A} \otimes \boldsymbol{e}_{0B} \qquad \boldsymbol{e}_{3A} \otimes \boldsymbol{e}_{1B}$$
$$\boldsymbol{e}_{3A} \otimes \boldsymbol{e}_{2B} \qquad \boldsymbol{e}_{3A} \otimes \boldsymbol{e}_{3B}$$

Using the pairs we can construct various mixtures in the same way we did it with a single register. For example, we could take 30% of $e_{1A} \otimes e_{3B}$, 25% of $e_{3A} \otimes e_{0B}$, 25% of $e_{2A} \otimes e_{1B}$ and 20% of $e_{0A} \otimes e_{2B}$: *Mixing combined register systems*

$$p = 0.3\, e_{1A} \otimes e_{3B} + 0.25\, e_{3A} \otimes e_{0B}$$
$$+\, 0.25\, e_{2A} \otimes e_{1B} + 0.2\, e_{0A} \otimes e_{2B}$$

and this mixture would still leave 12 other basis states unused.

If ω_A is a form (or a measurement) that acts on states of register A and η_B *Forms on* is a form (or a measurement) that acts on states of register B, then the two can *combined* be combined into a form $\omega_A \otimes \eta_B$ that acts on states of the combined register as *register states* shown in the following example:

$$\langle \omega_A \otimes \eta_B, p \rangle$$
$$= \langle \omega_A \otimes \eta_B, 0.3\, e_{1A} \otimes e_{3B} + 0.25\, e_{3A} \otimes e_{0B}$$
$$+\, 0.25\, e_{2A} \otimes e_{1B} + 0.2\, e_{0A} \otimes e_{2B} \rangle$$
$$= 0.3\langle \omega_A \otimes \eta_B, e_{1A} \otimes e_{3B} \rangle + 0.25\langle \omega_A \otimes \eta_B, e_{3A} \otimes e_{0B} \rangle$$
$$+\, 0.25\langle \omega_A \otimes \eta_B, e_{2A} \otimes e_{1B} \rangle + 0.2\langle \omega_A \otimes \eta_B, e_{0A} \otimes e_{2B} \rangle$$
$$= \ldots$$

Now we proceed exactly as we did in the definition of the operator. Let us first unite A forms with A vectors. This will produce numbers that will get thrown out in front of the $\langle$ and $\rangle$ brackets leaving B forms and B vectors to do the same. So in final account we shall get

$$\ldots = 0.3\langle \omega_A, e_{1A} \rangle \langle \eta_B, e_{3B} \rangle + 0.25\langle \omega_A, e_{3A} \rangle \langle \eta_B, e_{0B} \rangle$$
$$+\, 0.25\langle \omega_A, e_{2A} \rangle \langle \eta_B, e_{1B} \rangle + 0.2\langle \omega_A, e_{0A} \rangle \langle \eta_B, e_{2B} \rangle$$

A function so defined is clearly linear on states $p \in \mathbb{R}^{16}$ and so it can be interpreted as arithmetic mean average over the statistical ensemble that corresponds to p.

Because under the action of $\omega_A \otimes \eta_B$ a pair such as $e_{0A} \otimes e_{3B}$ gets converted into *A formal (or* a *product* of two reals, $\langle \omega_A, e_{0A} \rangle \langle \eta_B, e_{3B} \rangle$, one can think of the tensor product $\otimes$ *tensor) product* as a *product in waiting*. *of register states*

Another way to think of a tensor product is as a logical *and*. Instead of saying that we have observed the pair of registers A and B in state

$$e_{0A} \otimes e_{3B}$$

we can say that we have observed the pair in state:

$$\boldsymbol{e}_{0A} \quad \text{and} \quad \boldsymbol{e}_{3B}$$

2 The Qubit

2.1 The ugly quanta

Macroscopic matter, i.e., things that surround us in our every day life, like cups, saucers, telephones and frog infested ponds, are subject to well known and well understood laws of macroscopic physics. The laws are pretty simple, although when applied to almost any realistic system they tend to yield very complex equations that are almost impossible to understand, seldom admit analytical solutions and frequently display chaotic behavior. It is often the case that we have to help ourselves with common-sense understanding of macroscopic world in order to construct, analyze and solve equations that describe it.

Deceptive simplicity of macroscopic physics

XIXth century physicists expected that laws of macroscopic physics, which they distilled from their macroscopic observations and refined with their macroscopic brains should extend do microscopic domain as well. They imagined that atoms (and they did not suspect at the time that atoms could be made of even smaller constituents) would be subject to the same laws of Newtonian dynamics that worked so well on cannon balls and anvils sliding off rotating wedges with rough surfaces.

It is just as well that they were wrong, because if they weren't, nothing would work and we could not and would not be here to discuss these issues. Matter based on the principles of classical physics would cease to exist almost instantaneously. The Rutherford model predicted that a typical lifetime of a hydrogen atom should be about 10^{-10} s. Even if we were to ignore this little difficulty, classical stars should run out of puff in a mere one hundred million years, as Kelvin and Helmholtz discovered. Needless to say, Rutherford knew very well that hydrogen atoms did not decay after only 10^{-10} s and Kelvin and Helmholtz were well aware of the fact that our Mother Earth was several billion years old.

According to classical physics we should not be here.

Physics of microscopic domain is then quite different, and one should indeed expect this. A typical macroscopic chunk of matter contains about the Avogadro number of molecules. It may be $1/100^{\text{th}}$ or $1/1000^{\text{th}}$ or perhaps a thousand times more than the Avogadro number. It does not make much difference, because the Avogadro number is so huge: 6.023×10^{23}/mole. The laws of microscopic physics are not only very different, they also appear richer than the laws of macroscopic physics. But when you put 6.023×10^{23} quantum objects together, couple them to other equally voluminous lumps of macroscopic matter and immerse the whole lot in a thermal bath of the Avogadro number of photons, the spectrum of which corresponds to the room temperature, most of this different and rich microscopic

Microscopic physics averages away in thermodynamic limit.

physics is going to *average away* and all we'll be left with in the macroscopic world will be our well known and intuitively sound macroscopic physics... which predicts that we should not exist.

One should expect that macroscopic physics ought to be derivable from microscopic physics. How would one go about this? One would derive mathematical formulae describing the behavior of a system of N quantum objects interacting with each other. This may not be possible to do exactly because of the complexity involved, but one could make some simplifying assumptions on the way. One would then take a limit $N \to \infty$ (but in such a way that $N/V = \text{density} = \text{constant}$) since the Avogadro number is large enough to be replaced by infinity, and one would also assume ambient temperature to be sufficiently high for quantum statistics to be replaced by the Boltzmann statistics, and in this limit we would expect the laws of macroscopic physics to emerge. Such a procedure is called *taking thermodynamic limit of a quantum theory*. It is indeed the case that laws of microscopic physics, as we know and understand them today, yield laws of macroscopic physics in thermodynamic limit.

Microscopic physics cannot be derived from macroscopic physics.

One should *not* expect that *microscopic* physics could be derived from macroscopic physics. Because of the *averaging away* of quantum effects on taking the thermodynamic limit, various theories of microscopic physics, some of them blatantly at odds with each other and with experimental phenomenology as well, may yield the same macroscopic physics in thermodynamic limit. A simple example is the Boltzmann statistics, which can be derived both from the Fermi-Dirac statistics and from the Bose-Einstein statistics, the latter two being quite different. There is no way then that, say, the Fermi-Dirac statistics can be derived from the Boltzmann statistics.

Smuggling macroscopic concepts into the domain of microscopic physics

One should neither expect that macroscopic concepts such as space-time continuum or differential manifolds would be applicable to the description of microscopic systems. There is no reason to expect that microscopic "space" should even be Hausdorff or, indeed, topological[1]. Yet one has to start somewhere and the devil we know is better than the one we don't. So in effect physicists have *smuggled* a lot of macroscopic conceptual framework into their description of microscopic world

---

[1]Consider the example of neutron decay. Free neutrons live 885.7 ± 0.8 seconds on average [35]. Their lifetime is measured with macroscopic clocks and we see some neutrons living less than 885.7 seconds and some more. The decay rate is exponential and the process of decay is probabilistic. But what is the macroscopic time that is used in the measurement? What if every neutron has its own microscopic clock and all these clocks work at different speeds, the macroscopic time being an average of all the little microscopic times? The neutron's lifetime measured by its own clock could very well be fixed, but because their clocks work at different speeds, we would perceive the neutron decay process as random when measured against the macroscopic clock.

and whereas it seems to work in general one is tempted to wonder sometimes if use of such concepts in microscopic physics is not *abuse*.

How then can we arrive at correct theories of microscopic world? The answer is simple if disappointing: educated guesswork combined with laboratory verification. The so called *quantization procedures* are just educated guesswork. They are not real derivations and they seldom yield correct quantum theories without the need for additional man-handling anyway. Their real purpose is to ensure that whatever is eventually concocted on the microscopic level yields an expected macroscopic theory in thermodynamic limit, and this is fair enough. It does not imply, however, that microscopic theories cannot be constructed in other ways, while still yielding correct thermodynamic limit. And, indeed, there are a good few ways to choose from, the most popular being canonical quantization, Feynman path integrals, spin networks, and more recently topological field theories[2].

Scientific method and microscopic physics

But at the end of the day all that matters is the microscopic description itself, which may as well have been guessed, its predictions and its laboratory verification. And the latter turns out to be a real can of worms because, as yet, nobody has come up with a way to observe a microscopic system without entangling it at the same time with a macroscopic (or at best a *meso*scopic) measuring apparatus. So, in effect, we don't really know what microscopic systems do when they are left on their own and how they interact with each other in absence of the macroscopic measuring apparatus – although we may have some vague ideas about it. The presence of the measuring apparatus in our investigations of microscopic world is so important that we have been forced to acknowledge that certain measured quantities and perhaps even measured microscopic objects themselves are *made* by the act of the measurement and they do not exist in the same form prior to the measurement.

Microscopic measurement is extremely difficult

But there is no need to panic. There is nothing sacred and incomprehensible about the measurement and the issues of conscience, perception and religious beliefs are not a part of it. Measurement is a dynamic physical process that represents interaction of a microscopic system with a system comprising the Avogadro number of other microscopic systems. It is a lopsided process but it can be understood, analyzed and verified even within the existing framework of microscopic physics.

Measurement is a physical process that can be analyzed within the domain of microscopic physics.

---

[2]I have once met a scientist who so believed in the physical reality of quantization procedures that she thought of a special device based on a quantization procedure she worked on. The purpose of the device was to manipulate quantum systems in a special way. She was most surprised to learn that the device could not be made, because her quantization procedure did not and could not correspond to any real physical process. The quantization procedure was merely a mathematical mapping between certain macroscopic concepts and certain quantum properties, similar to the mapping between Poisson brackets and commutators.

This is indeed one of physics' greatest accomplishments in the last 20 years of the XXth century.

The difficulties of microscopic measurement can be understood better by pondering what it means *to observe*.

We observe behavior of macroscopic objects by *looking* at them or looking at instruments that, in turn, look at the objects. "Looking" at something implies that we have a light source emitting photons, which bounce off the observed object, then enter our eyes and get absorbed by the retina, which converts light into chemical energy. The chemical energy activates nerves that transmit the signals to the brain, which interprets them. The crucial link in this chain is where photons bounce off the observed object. The amount of momentum transmitted from the photons to the object is so small compared to the momentum of the object that the act of illumination does not affect the object and its behavior appreciably. Consequently we can ignore the effect that the light source has on the object. The planets do not change their orbits and rotation because the sun shines on them. The anvil sliding off the rotating wedge with the rough surface is not going to stop suddenly and then jump back to the top, because a laboratory lamp has been turned on[3].

Fragility of microscopic systems

But suppose that instead of using a benign light source in the form of a candle or a light bulb we were to use an exploding nuclear bomb. This would certainly be an overkill and we would end up none the wiser, because the explosion would obliterate anything we would try to look at. Yet this is very much what happens in the microscopic world. Microscopic systems are so delicate, so fragile, that even bouncing as little as a single photon against them can change their state dramatically. This state of affairs is exacerbated further by the fact that the wavelength of a photon is inversely proportional to its momentum. A photon of low momentum has large wavelength, so it is not going to be a precise enough instrument with which to observe microscopic systems. But if we attempt to select a photon with wavelength sufficiently short to give us a well resolved picture of a microscopic object, its momentum will be so large that it will destroy the object we are trying to observe. Because the same relation affects all other elementary particles, we can't eliminate this difficulty by choosing, e.g., neutrons to observe microscopic systems – although it is sometimes possible to get just a little further by observing with particles other than photons.

The Heisenberg Uncertainty Principle

Non-locality of microscopic objects

But there is a yet another complication which is perhaps the weirdest of all manifestations of quantum physics. Quantum objects such as photons and electrons

---

[3]Though a photographic emulsion, of course, is going to react with the incident light, unless it is of a wavelength to which the emulsion is insensitive – but here we are almost at the border of quantum physics, so let's keep this particular example in the footnote.

are non-local. They only turn into point-like energy discharges when they get snatched away from their free-range status by a macroscopic measuring apparatus. It is as if the apparatus sucked them into the point. All other free-ranging quantum systems the measured particle interacted with prior to this act of kidnaping and localization detect the sudden absence of their companion and react to it in various ways. So in quantum physics it is not just the act of "shining light" onto an observed object that affects it. The act of kidnaping bounced-off photons affects the object too. It is as if the anvil sliding off a rotating wedge with a rough surface stopped suddenly and jumped back to the top because we gave it a furtive glance.

Einstein, the great physicist of XXth century, could not stomach this. Yet stomach this he had to, because this is what clearly transpired from laboratory experiments. He responded by writing about "the ugly quanta" in some of his last letters. But quanta aren't ugly. They are what they are and the best way to make them likeable is to understand their weirdness and seek to explore it. *Einstein's curse*

So, what can we say about microscopic systems? After all we have been investigating them for some 100 years or so. Numerous Nobel prizes were awarded for successful predictions pertaining to and then discoveries and exploitations of various microscopic phenomena. Devices such as lasers and semiconductor switches, the functioning of which is based on principles of quantum physics, are incorporated into common household appliances.

The first thing to observe is that microscopic systems are essentially unpredictable. It is usually impossible to predict exactly what a given single electron or a photon is going to do in various experimental contexts, although there are some rare situations for which such predictions can be made. When a single microscopic system is subjected to a "nuclear blast" of a measurement with a macroscopic apparatus we end up with random read-outs, which, to make things worse, may not always tell us what the microscopic system's properties were prior to the measurement. As we have already pointed out, some properties and perhaps even the objects themselves appear to be made by the measurement. *Unpredictability of individual microscopic systems*

But this random read-out is not necessarily white-random. If we repeat the measurement over and over on microscopic systems that have been prepared in exactly the same way, we'll discover that there may be certain probability distributions associated with experimental results. And these distributions, it turns out, can be predicted with great accuracy. *Probabilistic analysis of microscopic systems*

Quantum physics is a discipline that tells us how to describe, manipulate and predict evolution of probability distributions associated with measurements made on microscopic systems with macroscopic measuring devices. *Quantum physics*

Much has been said in the past about universality of quantum physics (see, e.g., *Universality of quantum physics*

[44] and references therein). At first glance quantum physics cannot be universal, because it describes interaction between microscopic systems and macroscopic devices (the measuring apparatus or just a macroscopic environment – as far as the microscopic system is concerned it's the same thing). But microscopic systems are *always* in the presence of some macroscopic environment so the question about what microscopic systems may possibly do "on their own" may perhaps be just as silly as the question about the number of angels that can dance on the tip of a needle.

Frontiers of quantum physics   And then again, once physicists have a theory such as quantum physics in hand, they always try to extend it and apply to phenomena that at first glance may be beyond the theory's original area of competence. This is a very worthwhile endeavor, because it is through such activities that our knowledge and understanding of nature expands too. And so, one can try to observe microscopic systems with *meso*scopic devices – the present day technology allows, at last, for such measurements to be carried out [40] [28] [47] – and one can then test, whether predictions of quantum physics still agree with what such mesoscopic measurements return. What we find then is that certain older formulations of quantum mechanics may have to be ever so subtly revised and enriched in order to account for what the new experiments tell us.

Quantum physics and quantum computing   These activities are also of great importance to quantum computing, because it is by these means that we learn how to manipulate microscopic systems to our advantage. And it is exactly here that the frontiers of present day quantum physics and computer science meet.

This is going to be our battlefield.

2.2   The fiducial vector of a qubit

What is a qubit?   A qubit is a quantum relative of a classical randomly fluctuating one-bit register. It is the simplest non-trivial quantum system.

How can qubits be made?   As is the case with randomly fluctuating one-bit registers, there are many possible physical embodiments of a qubit. Qubits can be "natural", e.g., a neutron placed in a very strong uniform magnetic field of about $12\,\mathrm{T}$ is an example of a natural qubit – and we'll work with this example a lot. Qubits can be engineered too, e.g., the so called *quantronium* circuit [52] makes an excellent qubit.

Qubits can be described in terms of fiducial vectors.   Yet, regardless of the details of their engineering, which can be sometimes quite complex, qubits' fiducial mathematics and dynamics are always the same. Qubits can be described in terms of fiducial vectors indeed. Each entry in the vector specifies a probability of finding a qubit in the corresponding configuration.

A qubit is a two-dimensional system ($N = 2$) in the sense that has been explained *The basis states* in section 1.5 on page 19. This means that if we were to give a qubit a quick glance, *of a qubit* speaking figuratively, of course, since glancing at quantum objects is far from trivial, we would find it in one of two possible states. These states are often referred to as "up" and "down" and denoted by symbols

$$|\uparrow\rangle \quad \text{and} \quad |\downarrow\rangle$$

We may associate the binary number 0 with $|\uparrow\rangle$ (because $|\uparrow\rangle$ is often the lower *Qubit computing* energy state of the two) and the binary number 1 with $|\downarrow\rangle$ and we end up with an object that, like a bit, can be used for counting. And this leads to the following notation that has been adopted in quantum computing:

$$|\uparrow\rangle \equiv |0\rangle$$
$$|\downarrow\rangle \equiv |1\rangle$$

A fiducial vector that describes a qubit is four-dimensional ($K = 4$) and can be *The fiducial* parameterized as follows: *vector of a qubit*

$$\boldsymbol{p} = \frac{1}{2} \begin{pmatrix} 1 + r^z \\ 1 - r^z \\ 1 + r^x \\ 1 + r^y \end{pmatrix} \tag{2.1}$$

where

$$(r^x)^2 + (r^y)^2 + (r^z)^2 \leq 1 \tag{2.2}$$

This fiducial vector very succinctly illustrates several fundamental differences between classical and quantum systems. First, even though the system is two-dimensional, its fiducial vector is four-dimensional ($K = 4 = N^2$) [26]. We will no longer have the trivial mapping between basis states and fiducial states that characterized classical randomly fluctuating registers. In the quantum world fiducial level description is quite different from the basis state level description.

The second thing to observe is that

$$\sum_{i=0}^{3} p^i = 2 + \frac{r^x + r^y}{2}$$

Vector $\boldsymbol{p}$ is normalized, but not in the classical sense. The normalization is restricted to its first two components only, i.e.,

$$p^0 + p^1 = 1$$

In order to explain what this means we have to focus on a specific physical realization of a qubit and describe how probabilities p^0 through p^3 can be measured.

Magnetic properties of the neutron

Perhaps the simplest such realization is a beam of neutrons. Although neutrons don't have electric charge, they are known to have magnetic moment [1]

$$\mu_n = 1.913148 \pm 0.000066 \, \mu_N$$

where

$$\mu_N = \frac{q_e \hbar}{2m_p} = 5.051 \times 10^{27} \, \text{Am}^2 \tag{2.3}$$

is the nuclear magneton and where q_e is the elementary charge, m_p is the mass of the proton, and $\hbar$ is the Planck constant divided by 2π.

Spatially varying magnetic field exerts a force on the neutron.

According to classical physics the mechanical energy of the neutron immersed█ in magnetic field $\boldsymbol{B}$ is[4]

$$E = -\boldsymbol{\mu}_n \cdot \boldsymbol{B} \tag{2.4}$$

If the magnetic field varies in space then there is a force exerted on the neutron by the gradient of the magnetic field $\boldsymbol{\nabla} \boldsymbol{B}$:

$$\boldsymbol{F} = -\boldsymbol{\nabla} E = \boldsymbol{\nabla} \left(\boldsymbol{\mu}_n \cdot \boldsymbol{B} \right) \tag{2.5}$$

Suppose that $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$. For the magnetic field so defined $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ and $\boldsymbol{\nabla} \times \boldsymbol{B} = -\alpha \boldsymbol{e}_y$. $\boldsymbol{B}$ has vanishing divergence, as it should, and it can be generated by current density $\boldsymbol{j} = -\epsilon_0 c^2 \alpha \boldsymbol{e}_y$.[5] To be more precise, the Maxwell equations tell us that there ought to be a B^x component varying with z in this situation too. The field has the configuration as shown in figure 2.1.

The solution $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$ corresponds to $\boldsymbol{B}$ along the x axis. We will confine ourselves to the narrow neighborhood of the x axis then in order to have $\boldsymbol{B}$ described by this formula.

The mechanical energy of the neutron immersed in $\boldsymbol{B}$ is $-\mu_n^z \alpha x$ and the force acting on it becomes

$$\boldsymbol{F} = \mu_n^z \alpha \boldsymbol{e}_x \tag{2.6}$$

Assuming that the neutron's magnetic moment can point in every direction, $-\mu_n \leq \mu_n^z \leq \mu_n$, which, in turn, yields force $\boldsymbol{F} = F^x \boldsymbol{e}_x$, where

$$-\mu_n \alpha \leq F^x \leq \mu_n \alpha \tag{2.7}$$

Generation of the magnetic field with required properties.

Suppose that a well collimated monochromatic and unpolarized[6] beam of neutrons is fired in the y direction as shown in figure 2.2. Then the beam enters a chamber filled with $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$. On leaving the chamber the beam should fan out in the x direction.

---

[4]To brush up on the dynamics of a current loop in the magnetic field see, e.g., [21], section 15-1, "The forces on a current loop; energy of a dipole".

[5]To brush up on the Maxwell equations see, e.g., [21], chapter 18, "The Maxwell Equations".

[6]*Collimated* means *well focused*, *monochromatic* means that all neutrons have the same energy and momentum, *unpolarized* – well, this will be explained later; just read on...
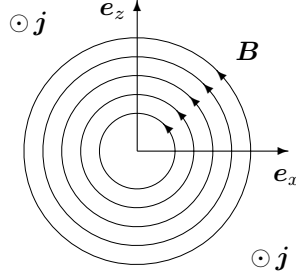
Figure 2.1: The magnetic field configuration for $\boldsymbol{j} = -\epsilon_0 c^2 \alpha \boldsymbol{e}_y$. The symbol $\odot$ to the left of $\boldsymbol{j}$ stands for the tip of the arrow that points at the reader. We find that $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$ along the x axis.
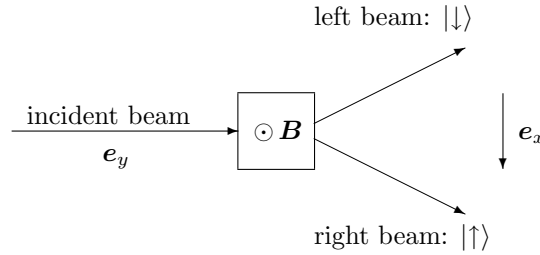
Figure 2.2: Splitting of the incident neutron beam by the chamber filled with $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$. The symbol $\odot$ inside the box representing the chamber stands for the tip of the arrow that points at the reader.

This is what classical physics says but this is not what happens.

Instead, the beam *splits*, as shown in figure 2.2, into two well collimated beams, one of which corresponds to $F^x = -\mu_n \alpha$ and the other one to $F^x = \mu_n \alpha$, and the whole middle that corresponds to $-\mu_n \alpha < F^x < \mu_n \alpha$ is missing. This result tells us that as the neutrons encounter $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$ they either "align" or "counter-align" with the direction of the ambient magnetic field. Furthermore, the "alignment" appears to happen instantaneously and in such a way that no energy is released in the process. Nobody has ever managed to capture a neutron that would be inclined under some angle to the direction of the ambient magnetic field and that would then gradually "align" with it, releasing the excess of $-\boldsymbol{\mu}_n \cdot \boldsymbol{B}$ in the process.

A classical physics phenomenon that resembles closely what we observe here is a passage of a light beam through a birefringent crystal. The crystal

The beam splits instead of fanning out

The splitting of the neutron beam resembles the splitting of a light beam by a birefringent crystal.

splits the beam into two components, which become physically separated from each other. Each component is linearly polarized in a direction perpendicular to that of the other component.[7]

Drawing on this similarity we refer to the two neutron beams that emerge from our apparatus as *fully polarized*. The beam that corresponds to $F^x = \mu_n \alpha$ will be deflected to the right[8] and the beam that corresponds to $F^x = -\mu_n \alpha$ will be deflected to the left. All neutrons in the right beam have their $\boldsymbol{\mu}_n$ aligned with $\boldsymbol{B}$ and all neutrons in the left beam have their $\boldsymbol{\mu}_n$ aligned against the direction of $\boldsymbol{B}$. Let us call the state of the ones in the right beam $|\uparrow\rangle$ and the state of the ones in the left beam $|\downarrow\rangle$.

Normalization of the fiducial vector

Now we can ask questions about neutrons comprising the incident beam. We can ask, for example, about the probability that a neutron entering the chamber will emerge from it in state $|\uparrow\rangle$. This probability is p^0. We can then ask about the probability that a neutron entering the chamber will emerge from it in state $|\downarrow\rangle$. This probability is p^1. Because the chamber does not swallow neutrons, every neutron should emerge from it in either of the two possible states. Consequently the probabilities p^0 and p^1 must add to 1:

$$p^0 + p^1 = 1$$

This is accomplished automatically by the parameterization:

$$
\begin{aligned}
p^0 &= \frac{1}{2}\left(1 + r^z\right) \\
p^1 &= \frac{1}{2}\left(1 - r^z\right)
\end{aligned}
$$

where $-1 \leq r^z \leq 1$.

Rotating the chamber by $90°$ *lets us measure* p^2.

If we were to rotate the whole set up about the y axis by $90°$ clockwise, we would end up with the chamber filled with $\boldsymbol{B} = \beta z \boldsymbol{e}_x$. Trying again the classical description, the energy of a neutron entering the chamber would be $E = -\mu_n^x \beta z$ and the force exerted on the neutron would be $\boldsymbol{F} = -\boldsymbol{\nabla} E = \mu_n^x \beta \boldsymbol{e}_z$. This time the beam of incident neutrons should fan out in the z direction. And this is again wrong, because the beam will instead *split* in the z direction, the two new beams corresponding to $F^z = \mu_n \beta$ and $F^z = -\mu_n \beta$ with the whole middle $-\mu_n \beta < F^z < \mu_n \beta$ missing. This is shown in figure 2.3.

Neutrons in the upper beam have their magnetic moments $\boldsymbol{\mu}_n$ aligned with the direction of the magnetic field $\boldsymbol{B} = \beta z \boldsymbol{e}_x$, whereas neutrons in the lower beam have their magnetic moments aligned against the direction of $\boldsymbol{B}$. Let us call the states of the neutrons in the upper beam $|\rightarrow\rangle$ and the states of the neutrons in the lower beam $|\leftarrow\rangle$. We shall call the probability that a

---

[7]Birefringence is discussed in [19], section 33-3, "Birefringence".

[8]Here we assume that the three vectors $\boldsymbol{e}_x$, $\boldsymbol{e}_y$ and $\boldsymbol{e}_z$ have the right hand screw orientation, i.e., as you rotate from $\boldsymbol{e}_x$ to $\boldsymbol{e}_y$ you move up in the direction of $\boldsymbol{e}_z$. We also assume that $\alpha > 0$.

upper beam: $|\rightarrow\rangle$

incident beam

$\boldsymbol{e}_y$

$\odot \boldsymbol{B}$

$\boldsymbol{e}_z$

lower beam: $|\leftarrow\rangle$
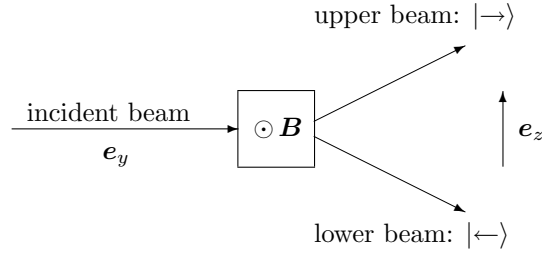
Figure 2.3: Splitting of the incident neutron beam by the chamber filled with $\boldsymbol{B} = \beta z \boldsymbol{e}_x$. As before, the symbol $\odot$ inside the box representing the chamber stands for the tip of the arrow that points at the reader.

neutron incident on the chamber emerges from it in the upper beam p^2 and we'll parameterize it by

$$p^2 = \frac{1}{2}\left(1 + r^x\right)$$

where $-1 \leq r^x \leq 1$. The probability that the neutron emerges from the chamber in the lower beam is then $1 - p^2 = \frac{1}{2}\left(1 - r^x\right)$, because every neutron that enters the chamber *must* leave it either in the upper or in the lower beam.

Finally, consider shooting the beam in the y direction into the chamber pervaded by $\boldsymbol{B} = \gamma x \boldsymbol{e}_y$. Neutrons entering the chamber acquire energy $E = -\boldsymbol{\mu}_n \cdot \boldsymbol{B} = -\mu_n^y \gamma x$ and they will be subjected to force $\boldsymbol{F} = -\boldsymbol{\nabla} E = \mu_n^y \gamma \boldsymbol{e}_x$. Classically, on leaving the chamber the beam should fan out in the x direction and, again, this is not what is going to happen. The beam will *split* in the x direction producing two new beams, which will correspond to $\boldsymbol{\mu}_n$ aligned with $\boldsymbol{B} = \gamma x \boldsymbol{e}_y$ (the right beam) and aligned against the direction of $\boldsymbol{B}$ (the left beam). This is shown in figure 2.4

Measuring p^3.

Let us call the state of neutrons in the right beam, i.e., the ones aligned with $\boldsymbol{B}$, $|\otimes\rangle$ and the state of neutrons in the left beam, i.e., the ones aligned against the direction of $\boldsymbol{B}$, $|\odot\rangle$. We have used this notation already in figures 2.2 and 2.3. Symbol $\otimes$ is suggestive of an arrow flying away from the reader – in this case the arrow points in the direction of the beam itself. Symbol $\odot$ is suggestive of an arrow flying towards the reader – in this case the arrow points against the direction of the beam.

We can ask about the probability that a neutron entering the chamber will leave in the $|\otimes\rangle$ state. This probability is p^3 and we can parameterize it by

$$p^3 = \frac{1}{2}\left(1 + r^y\right)$$

where $-1 \leq r^y \leq 1$. The probability of finding the neutron in the $| \odot \rangle$ state will be $1 - p^3 = \frac{1}{2}(1 - r^y)$, because every neutron that enters the chamber *must* exit it either in the $| \otimes \rangle$ or in the $| \odot \rangle$ state.

Probabilities p^0, p^2 and p^3 cannot be measured at the same time.
Observe that we cannot measure simultaneously p^0, p^2 and p^3. The apparatus needed to measure p^0 is oriented differently from the apparatus needed to measure p^2 or p^3. In order to measure all three quantities we have to subject the beam of identically prepared neutrons to all three measurements separately.
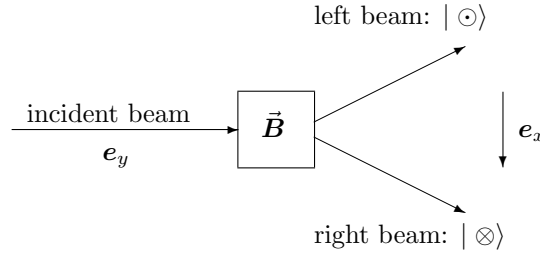
Figure 2.4: Splitting of the incident neutron beam by the chamber filled with $\boldsymbol{B} = \gamma x \boldsymbol{e}_y$. The arrow above $\boldsymbol{B}$ indicates that the magnetic field inside the chamber points to the right.

Probabilities p^0, p^2 and p^3 do not describe mutually exclusive configurations.
The measurements do not represent mutually exclusive alternatives either. If we were to produce a polarized beam in the $| \uparrow \rangle$ state and then pass it through a device that should split it between $| \rightarrow \rangle$ and $| \leftarrow \rangle$ states, the beam would indeed split and it would split, in this case, evenly. This is why the components of the fiducial vector $\boldsymbol{p}$ add to more than 1. Yet, in order to fully characterize the state of a qubit, we have to measure all three probabilities: p^0, p^2 and p^3.

2.3   Polarized states

Now that we know how to measure probabilities that form the fiducial vector $\boldsymbol{p}$ of a qubit, let us consider certain special situations. We shall focus again on the neutron beam embodiment of a qubit.

Beams leaving the beam splitting apparatus are fully polarized
Suppose the beam is fired in the y direction whereupon it enters a chamber filled with $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$. The beam will split in the x direction so that neutrons in state $| \uparrow \rangle$ will shoot to the right and neutrons in state $| \downarrow \rangle$ will shoot to the left as shown

in figure 2.2. The fiducial vector that describes neutrons in the right beam looks as follows:

$$|\uparrow\rangle \equiv \begin{pmatrix} 1 \\ 0 \\ 0.5 \\ 0.5 \end{pmatrix}$$

Vector $\boldsymbol{r}$ that sits inside $\boldsymbol{p}$ points upwards

$$\boldsymbol{r} = \boldsymbol{e}_z$$

The fiducial vector that describes neutrons in the left beam is

$$|\downarrow\rangle \equiv \begin{pmatrix} 0 \\ 1 \\ 0.5 \\ 0.5 \end{pmatrix}$$

Vector $\boldsymbol{r}$ that sits inside $\boldsymbol{p}$ points downwards this time

$$\boldsymbol{r} = -\boldsymbol{e}_z$$

Observe that p^2 and p^3 in both fiducial vectors are not zero. This is because in both cases we have that $r^x = r^y = 0$, but this merely leaves $p^2 = p^3 = \frac{1}{2}(1 + 0) = \frac{1}{2}$. What does this mean?

Let us take the right beam (neutrons in state $|\uparrow\rangle$) and let us direct it into an- *A fully polarized* other chamber filled with $\boldsymbol{B} = \beta z \boldsymbol{e}_x$ as shown in figure 2.3. The fiducial vector *beam is split* that corresponds to $|\uparrow\rangle$ tells us that the beam will split in this chamber evenly, *again.* i.e., approximately half of all neutrons (remember that probabilities become exact measures of what is going to happen only when the number of neutrons in the beam becomes infinite) that enter it will swing upwards and the other half will swing downwards. This prediction pertains to the statistical ensemble only. We have no means of predicting what any given neutron is going to do. It may just as well swing upwards as downwards, much the same as a well thrown coin can land just as well heads up or tails up. In any case, half of all neutrons, *on average*, will emerge from this chamber in the $|\rightarrow\rangle$ state and the other half will emerge in the $|\leftarrow\rangle$ state. Their corresponding fiducial vectors will be

$$|\rightarrow\rangle \equiv \begin{pmatrix} 0.5 \\ 0.5 \\ 1 \\ 0.5 \end{pmatrix} \quad \boldsymbol{r} = \boldsymbol{e}_x$$

and

$$|\leftarrow\rangle \equiv \begin{pmatrix} 0.5 \\ 0.5 \\ 0 \\ 0.5 \end{pmatrix} \qquad \boldsymbol{r} = -\boldsymbol{e}_x$$

The beam splitting apparatus destroys the state of the incident beam and replaces it with the mixture of output beams.

At first glance we might think then that $|\uparrow\rangle$ is a mixture of $|\rightarrow\rangle$ and $|\leftarrow\rangle$. But it is not so as the following calculation shows[9]

$$0.5 \begin{pmatrix} 0.5 \\ 0.5 \\ 1 \\ 0.5 \end{pmatrix} + 0.5 \begin{pmatrix} 0.5 \\ 0.5 \\ 0 \\ 0.5 \end{pmatrix} = 0.5 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

This fiducial vector corresponds to $\boldsymbol{r} = \boldsymbol{0}$, not to $\boldsymbol{r} = \boldsymbol{e}_z$. The state of the beam has been changed by interaction with the second chamber and whatever information was stored in the input state has been irretrievably lost. Whereas the input state was fully polarized, the output state contains no useful information: it is all white noise.

Superposition versus mixture

We will learn later that $|\uparrow\rangle$ is related to $|\rightarrow\rangle$ and $|\leftarrow\rangle$ indeed, but it is not a mixture of these two states. Instead we will find in chapter 4 that it is a superposition of these states. We will be able to write

$$|\uparrow\rangle = \frac{1}{\sqrt{2}} \left(|\rightarrow\rangle + |\leftarrow\rangle \right)$$

but this will *not* translate into

$$\boldsymbol{p}_{|\uparrow\rangle} = \frac{1}{2} \left(\boldsymbol{p}_{|\rightarrow\rangle} + \boldsymbol{p}_{|\leftarrow\rangle} \right)$$

As we have remarked earlier, the relationship between basis states and fiducial vectors in quantum physics is far from trivial.

If the beam splitting apparatus is set so as to confirm the state of the beam, the beam emerges from it intact.

Let us go back to the right beam that emerged from the first chamber, the beam that was in the $|\uparrow\rangle$ state. If instead of directing this beam into the chamber filled with $\boldsymbol{B} = \beta z \boldsymbol{e}_x$ we were to direct it into a chamber filled with $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$, i.e., a chamber that works exactly the same way as the first chamber, we would find that *all* neutrons entering the chamber would be swung to the right. This means that this time the chamber would not change their state. Instead the chamber would merely confirm the state of the incident neutrons. This is one of the rare circumstances in quantum physics when we can predict *with certainty* what is going to happen to *every* individual neutron.

---

[9]We're getting a little ahead here. Mixtures of qubit states will be discussed in section 2.4, page 56.

This behavior is very similar to the behavior of photons passing through a series of polarizing plates. Suppose the first plate polarizes incident light in the z direction. If we insert a second plate behind the first plate and rotate its polarization axis by $90°$ relative to the first plate, no photons will pass through the two plates. If we rotate the polarization axis of the second plate by $90°$ again, both plates will have their axes aligned and every photon that passes through the first plate is guaranteed to pass through the second plate as well. We can make this prediction *with certainty* about every photon incident on the second plate. But if the second plate has been rotated by $45°$ relative to the first plate, only about a half of the photons incident on the second plate will be transmitted. Yet, we have no means to make exact predictions about any individual photon incident on the second plate in this case. We can only make predictions about the statistical ensemble that, in this case, corresponds to the beam of the incident photons.

Neutron and photon polarization states are very similar.

So far we have contemplated polarized states aligned with one of the three principal directions, x, y, and z only. What about a state described by vector $\boldsymbol{r}$ of length 1 but tilted arbitrarily? We can describe components of such a vector by using spherical coordinates θ and ϕ (see figure 2.5)

Neutron beam polarized in an arbitrary direction

$$\boldsymbol{r} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}$$

θ is the angle between $\boldsymbol{e}_z$ and $\boldsymbol{r}$.

The corresponding fiducial vector is

ϕ is the angle between $\boldsymbol{e}_x$ and the projection of $\boldsymbol{r}$ on the equatorial plane, $z = 0$.

$$\boldsymbol{p} = \frac{1}{2} \begin{pmatrix} 1 + \cos\theta \\ 1 - \cos\theta \\ 1 + \sin\theta\cos\phi \\ 1 + \sin\theta\sin\phi \end{pmatrix}$$

If we were to shoot a beam of neutrons so polarized in the y direction and through the chamber filled with magnetic field parallel to $\boldsymbol{r}$ and varying linearly in the direction perpendicular to $\boldsymbol{r}$ *all* incident neutrons would swing to the same side in the direction perpendicular to $\boldsymbol{r}$. If we can find such a direction by trial and error and confirm that all neutrons swing to the same side indeed, we'll know that the incident beam is fully polarized.

Procedure that tests whether the incident beam of neutrons is fully polarized

It is the same as with polarized light. If we have a fully polarized incident light beam and a polarizer, we will always be able to find an angle of the polarizer that lets all incident photons through. If, on the other hand, some photons get absorbed for *every* angle of the polarizer, the incident beam is not fully polarized. It is a *mixture* of photons polarized in various directions.

Procedure that tests whether the incident beam of photons is fully polarized

Fully polarized states of a neutron beam are described by vector $\boldsymbol{r}$ of length   *$\boldsymbol{r}$ of length 1*
1 pointing in some arbitrary direction. Such states are *pure*. They cannot be   *describes fully*
produced by mixing neutrons of various polarizations. We have seen what happened   *polarized states*
when we tried to mix neutrons in state $|\rightarrow\rangle$ with neutrons in state $|\leftarrow\rangle$.   *of neutron*

Bloch sphere   The set of all *pure* states of the neutron beam corresponds to the surface traced   *beams*
by vector $\boldsymbol{r}$ of length 1 as it points in all possible directions. This is the surface of
a sphere of radius 1. This sphere shown in figure 2.5 is called the *Bloch sphere*.
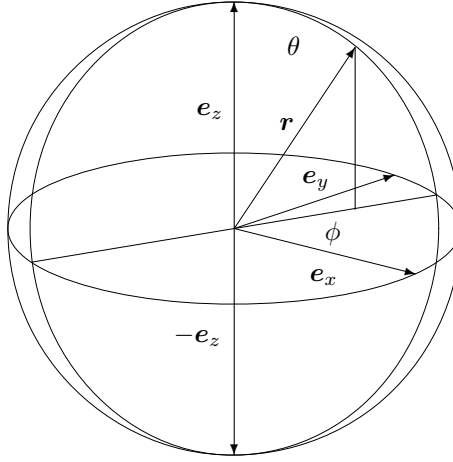
Figure 2.5: The Bloch sphere. The radius of the Bloch sphere is 1. All vectors
shown in the figure touch the surface of the Bloch sphere with their tips. θ is the
angle between $\boldsymbol{e}_z$ and $\boldsymbol{r}$ and ϕ is the angle between $\boldsymbol{e}_x$ and the projection of $\boldsymbol{r}$ on
the equatorial plane.

2.4   Mixtures of qubit states

Consider two qubit states. One given by

$$
\boldsymbol{p}_1 = \frac{1}{2}\begin{pmatrix} 1 + r_1^z \\ 1 - r_1^z \\ 1 + r_1^x \\ 1 + r_1^y \end{pmatrix}
$$

and the other one by

$$p_2 = \frac{1}{2} \begin{pmatrix} 1 + r_2^z \\ 1 - r_2^z \\ 1 + r_2^x \\ 1 + r_2^y \end{pmatrix}$$

We can construct a mixture of the two states in the same way we constructed mixtures for classical randomly fluctuating registers. The state of the mixture will be given by:

$$p = P_1 p_1 + P_2 p_2, \tag{2.8}$$

where $P_1 + P_2 = 1$. What is going to be the resulting vector r for the mixture?

Mixtures of qubit states are constructed the same way as mixtures of states describing classical randomly fluctuating registers.

$$p = \frac{P_1}{2} \begin{pmatrix} 1 + r_1^z \\ 1 - r_1^z \\ 1 + r_1^x \\ 1 + r_1^y \end{pmatrix} + \frac{P_2}{2} \begin{pmatrix} 1 + r_1^z \\ 1 - r_1^z \\ 1 + r_1^x \\ 1 + r_1^y \end{pmatrix} = \frac{1}{2} \begin{pmatrix} P_1 + P_2 + P_1 r_1^z + P_2 r_2^z \\ P_1 + P_2 - P_1 r_1^z - P_2 r_2^z \\ P_1 + P_2 + P_1 r_1^x + P_2 r_2^x \\ P_1 + P_2 + P_1 r_1^y + P_2 r_2^y \end{pmatrix}$$

We can now make use of $P_1 + P_2 = 1$ to collect this into:

$$p = \frac{1}{2} \begin{pmatrix} 1 + r^z \\ 1 - r^z \\ 1 + r^x \\ 1 + r^y \end{pmatrix}$$

where

$$r = P_1 r_1 + P_2 r_2$$

Suppose that both states that comprise the mixture, i.e., p_1 and p_2 correspond to fully polarized states, i.e., states for which $r_i \cdot r_i = 1$, where $i = 1$ or $i = 2$. Let us see now how mixing affects the length of r.

$$\begin{aligned} r \cdot r &= (P_1 r_1 + P_2 r_2) \cdot (P_1 r_1 + P_2 r_2) \\ &= \left(P_1^2 r_1^2 + 2P_1 P_2 r_1 \cdot r_2 + P_2^2 r_2^2 \right) \\ &= \left(P_1^2 + 2P_1 P_2 \cos\theta + P_2^2 \right) \end{aligned}$$

where θ is the angle between r_1 and r_2. When $\theta = 0$ we get

$$r \cdot r = \left(P_1^2 + 2P_1 P_2 + P_2^2 \right) = (P_1 + P_2)^2 = 1^2 = 1$$

but this is not a mixture, because $\theta = 0$ means that $p_1 = p_2$. For a real mixture $\theta \neq 0$, which means that $\cos\theta < 1$. In this case we get that

$$r \cdot r = \left(P_1^2 + 2P_1 P_2 \cos\theta + P_2^2 \right) < \left(P_1^2 + 2P_1 P_2 + P_2^2 \right) = 1$$

We see that mixing two fully polarized (and *different*) states results in a state with a shorter $\boldsymbol{r}$. *Mixing fully polarized qubit states produces a state with $\|\boldsymbol{r}\| < 1$.*

Bloch ball This means that states that fill the interior of the Bloch sphere are all mixtures. It also confirms that the fully polarized, pure states on the Bloch sphere cannot be produced by mixing other states, because mixing always delivers a state that is located somewhere inside the Bloch sphere. The pure states on the Bloch sphere are extremal, which is in agreement with our definition of pure states given on page 19. Set S of all possible physical mixtures and pure states of a qubit[10] forms a ball of radius 1. This ball is called the *Bloch ball*.

> Set S of all possible qubit states is quite different from the set of all possible states of a randomly fluctuating classical register. The latter is edgy with pure states well separated from each other. On the other hand pure, i.e., fully polarized states of a qubit form a smooth continuous surface, the Bloch sphere. The only reversible Cinderella transformations for a randomly fluctuating classical register were permutations of its pure states and the resulting transformations of mixtures. None of them were continuous. On the other hand, we have an infinite number of continuous reversible Cinderella transformations of a qubit, because a sphere can be mapped onto itself in an infinite number of rotations.

Consider a mixture state given by $\boldsymbol{r}$ where $\|\boldsymbol{r}\| < 1$. We can always rotate our system of coordinates to align z with $\boldsymbol{r}$ and we can always rotate the chamber that splits the incident neutron beam accordingly. Without loss of generality we can therefore assume that $\boldsymbol{r} = r\boldsymbol{e}_z$ and $0 \leq r < 1$. The fiducial vector that describes this mixture is

$$\boldsymbol{p} = \frac{1}{2} \begin{pmatrix} 1 + r \\ 1 - r \\ 1 \\ 1 \end{pmatrix}$$

A mixed beam splits for every possible direction of the magnetic field in the beam splitting chamber. Because $0 \leq r < 1$ we find that $0.5 \leq p^0 < 1$ and $0 < p^1 \leq 0.5$. This result tells us that for the mixture state the incident beam is *always* going to split, even if we happen to adjust magnetic field inside the chamber so that it is parallel to $\boldsymbol{r}$. There will be always some neutrons that will swing to the left, even if most (or at worst a half for $\boldsymbol{r} = \boldsymbol{0}$) will swing to the right. The beam is not fully polarized. On the

---

[10]Compare with set S for a classical randomly fluctuating register, figure 1.4 on page 17.

other hand, each of the two beams that leave the beam splitter, is fully polarized.
The one that swings to the right is described by

$$\boldsymbol{p}_{|\uparrow\rangle} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0.5 \\ 0.5 \end{pmatrix}$$

and the one that swings to the left is described by

$$\boldsymbol{p}_{|\downarrow\rangle} = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 0.5 \\ 0.5 \end{pmatrix}$$

2.5   The measurement

Consider a general fully polarized state given by

$$\boldsymbol{p} = \frac{1}{2}\begin{pmatrix} 1 + \cos\theta \\ 1 - \cos\theta \\ 1 + \sin\theta\cos\phi \\ 1 + \sin\theta\sin\phi \end{pmatrix}$$

Let us dwell a moment on what happens when this state is subjected to a full set
of measurements as described in section 2.2 on page 46.

Suppose the first measurement splits the beam in field $\boldsymbol{B} = \alpha x \boldsymbol{e}_z$. The measure- *Splitting a fully*
ment is going to align or counter-align each incident neutron with $\boldsymbol{e}_z$. Approxi- *polarized beam*
mately $(1 + \cos\theta)/2$ of all incident neutrons will swing to the right and will emerge *described by an*
from the apparatus in state   *arbitrary vector*
                                                                                                     r.

$$\boldsymbol{p}_{|\uparrow\rangle} = \begin{pmatrix} 1 \\ 0 \\ 0.5 \\ 0.5 \end{pmatrix}$$

The remaining neutrons, $(1 - \cos\theta)/2$ of the incident neutrons, will swing to the
left and emerge from the apparatus in state

$$\boldsymbol{p}_{|\downarrow\rangle} = \begin{pmatrix} 0 \\ 1 \\ 0.5 \\ 0.5 \end{pmatrix}$$

We can now merge both beams creating a new state:

$$\boldsymbol{p} \;=\; \frac{1}{2}\left(1+\cos\theta\right)\boldsymbol{p}_{|\uparrow\rangle} + \frac{1}{2}\left(1-\cos\theta\right)\boldsymbol{p}_{|\downarrow\rangle}$$

$$=\; \frac{1}{2}\begin{pmatrix} 1+\cos\theta \\ 1-\cos\theta \\ 1 \\ 1 \end{pmatrix}$$

The corresponding vector $\boldsymbol{r}$ is given by

$$\boldsymbol{r} = \begin{pmatrix} \cos\theta \\ 0 \\ 0 \end{pmatrix}$$

Merging the two beams that exit the beam splitter does not reconstruct the state of the incident beam. The length of this vector is 1 only for $\theta = 0$ which would mean that we have managed to align $\boldsymbol{B}$ inside the splitting apparatus with the original direction of $\boldsymbol{r}$. Otherwise $r^2 < 1$. The state created by merging beams that left the beam splitter is a mixture. We see here that the act of measurement destroys the original state and that state cannot be reconstructed by merely merging the two beams together. The act of measurement destroys information about the other angle, ϕ.

Wouldn't we be luckier then trying a measurement against $\boldsymbol{B} = \beta z \boldsymbol{e}_x$?

Let us try this option. The measuring apparatus in this case is going to split the beam so that $(1 + \sin\theta\cos\phi)/2$ of the incident neutrons will swing up. These neutrons will leave the apparatus in state $|\rightarrow\rangle$ described by

$$\boldsymbol{p}_{|\rightarrow\rangle} = \begin{pmatrix} 0.5 \\ 0.5 \\ 1 \\ 0.5 \end{pmatrix}$$

The remaining $(1 - \sin\theta\cos\phi)/2$ of the incident neutrons will swing down leaving the apparatus in state $|\leftarrow\rangle$ described by

$$\boldsymbol{p}_{|\leftarrow\rangle} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0 \\ 0.5 \end{pmatrix}$$

Merging the two beams together will produce the state described by

$$\boldsymbol{p} \;=\; \frac{1}{2}\left(1+\sin\theta\cos\phi\right)\boldsymbol{p}_{|\rightarrow\rangle} + \frac{1}{2}\left(1-\sin\theta\cos\phi\right)\boldsymbol{p}_{|\leftarrow\rangle}$$

$$= \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 + \sin\theta\cos\phi \\ 1 \end{pmatrix}$$

The corresponding vector $\boldsymbol{r}$ is given by

$$\boldsymbol{r} = \begin{pmatrix} 0 \\ \sin\theta\cos\phi \\ 0 \end{pmatrix}$$

Only for $\theta = 90°$ and $\phi = 0$ is $r^2 = 1$. For all other angles $r^2 < 1$ and so we end up with the mixture again.

Carrying out this measurement alone is not going to tell us much about either θ or ϕ. All we are going to see is some $r < 1$ and $\boldsymbol{r}$ pointing in the direction of $\boldsymbol{e}_x$. Great many combinations of θ and ϕ can produce such $\boldsymbol{r} = r\boldsymbol{e}_x$. But if we were to perform this measurement *after* the previous measurement on an identically prepared neutron beam, we would already have θ and this should give us $\cos\phi$. Yet, knowing $\cos\phi$ still does not let us determine $\boldsymbol{r}$ uniquely, because $\cos\phi = \cos(360° - \phi)$. We would have to find about the sign of $\sin\phi$ in order to determine ϕ uniquely. And it is here that the third measurement against $\boldsymbol{B} = \gamma x \boldsymbol{e}_y$ comes in. *Three measurements are required to determine ϕ and θ uniquely.*

Every measurement we have discussed in this section converts a pure state into a mixture, unless it happens to *confirm* the pure state. The resulting mixture contains less information than the original state. One has to carry out three independent measurements on the incident beam, ensuring that the state of the beam does not change between the measurements, in order to reconstruct the original state of the beam.

How would we go about measuring a mixture? The procedure discussed above refers to the fully polarized state, so we assume from the beginning that $r = 1$. We would still measure the three probabilities that would yield $r^x = r\sin\theta\cos\phi$, $r^y = r\sin\theta\sin\phi$ and $r^z = r\cos\theta$. Squaring and adding the three components will give us $(r^x)^2 + (r^y)^2 + (r^z)^2 = r^2$ and hence r. Now we can divide each of the three components by r and we end up with the same problem we had for the fully polarized state, i.e., we get θ from r^z/r, then $\cos\phi$ from r^x/r and finally we would get the sign of $\sin\phi$ from r^y/r in order to determine ϕ uniquely. *Measuring a mixture*

2.6 Pauli vectors and Pauli forms

Pauli vectors extracted from the fiducial vector of a qubit

The generic form of the qubit fiducial vector

$$\boldsymbol{p} = \frac{1}{2} \begin{pmatrix} 1 + r^z \\ 1 - r^z \\ 1 + r^x \\ 1 + r^y \end{pmatrix}$$

can be rewritten in the following way

$$\boldsymbol{p} = \frac{1}{2} \left(\boldsymbol{\varsigma}_1 + r^x \boldsymbol{\varsigma}_x + r^y \boldsymbol{\varsigma}_y + r^z \boldsymbol{\varsigma}_z \right) \tag{2.9}$$

where we are going to call

$$\boldsymbol{\varsigma}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \boldsymbol{\varsigma}_x = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad \boldsymbol{\varsigma}_y = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \qquad \boldsymbol{\varsigma}_z = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \tag{2.10}$$

Pauli vectors and Pauli matrices

Pauli vectors. Although Pauli didn't invent these four vectors, they are very closely related to Pauli matrices and fulfill a similar role to Pauli matrices within the fiducial formalism. We are going to use the wiggly symbol ς (pronounced "varsigma") for Pauli vectors bowing to tradition, because its close relative, the Greek letter $\boldsymbol{\sigma}$ (pronounced "sigma"), is used commonly to denote Pauli matrices.

Pauli vectors can be represented in terms of canonical basis vectors of the fiducial space

$$\boldsymbol{e}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{e}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

as follows:

$$\begin{aligned} \boldsymbol{\varsigma}_1 &= \boldsymbol{e}_0 + \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 \\ \boldsymbol{\varsigma}_x &= \boldsymbol{e}_2 \\ \boldsymbol{\varsigma}_y &= \boldsymbol{e}_3 \\ \boldsymbol{\varsigma}_z &= \boldsymbol{e}_0 - \boldsymbol{e}_1 \end{aligned} \tag{2.11}$$

This can be easily inverted to yield $\boldsymbol{e}_i$ in terms of $\boldsymbol{\varsigma}_i$:

$$\boldsymbol{e}_0 = \frac{1}{2} \left(\boldsymbol{\varsigma}_1 + \boldsymbol{\varsigma}_z - \boldsymbol{\varsigma}_x - \boldsymbol{\varsigma}_y \right)$$

$$
\begin{aligned}
\boldsymbol{e}_1 &= \frac{1}{2}\left(\boldsymbol{\varsigma}_1 - \boldsymbol{\varsigma}_z - \boldsymbol{\varsigma}_x - \boldsymbol{\varsigma}_y\right) \\
\boldsymbol{e}_2 &= \boldsymbol{\varsigma}_x \\
\boldsymbol{e}_3 &= \boldsymbol{\varsigma}_y
\end{aligned}
\tag{2.12}
$$

We are also going to introduce four Pauli forms defined by   *Pauli forms*

$$
\begin{aligned}
\boldsymbol{\varsigma}^1 &= (1,1,0,0) \\
\boldsymbol{\varsigma}^x &= (-1,-1,2,0) \\
\boldsymbol{\varsigma}^y &= (-1,-1,0,2) \\
\boldsymbol{\varsigma}^z &= (1,-1,0,0)
\end{aligned}
\tag{2.13}
$$

Pauli forms are dual to Pauli vectors, meaning that[11]   *Pauli vectors*
                                                                              and Pauli forms

$$
\langle \boldsymbol{\varsigma}^i, \boldsymbol{\varsigma}_j \rangle = 2\delta^i{}_j, \quad i,j = 1,x,y,z
\tag{2.14}
$$
                                                                              are mutually
                                                                              dual.

From this we can easily derive that

$$
\begin{aligned}
\langle \boldsymbol{\varsigma}^1, \boldsymbol{p} \rangle &= 1 \\
\langle \boldsymbol{\varsigma}^x, \boldsymbol{p} \rangle &= r^x \\
\langle \boldsymbol{\varsigma}^y, \boldsymbol{p} \rangle &= r^y \\
\langle \boldsymbol{\varsigma}^z, \boldsymbol{p} \rangle &= r^z
\end{aligned}
$$

We can therefore use Pauli forms as devices for extracting r^x, r^y and r^z from   *Pauli forms*
arbitrary fiducial vectors $\boldsymbol{p}$.   *extract $\boldsymbol{r}$ from*
    It is easy to express canonical forms   *$\boldsymbol{p}$.*

$$
\begin{aligned}
\boldsymbol{\omega}^0 &= (1,0,0,0) \\
\boldsymbol{\omega}^1 &= (0,1,0,0) \\
\boldsymbol{\omega}^2 &= (0,0,1,0) \\
\boldsymbol{\omega}^3 &= (0,0,0,1)
\end{aligned}
$$

in terms of Pauli forms and vice versa

$$
\boldsymbol{\omega}^0 = \frac{1}{2}\left(\boldsymbol{\varsigma}^1 + \boldsymbol{\varsigma}^z\right)
$$

---

[11]It would be quite possible to redefine Pauli vectors or forms so that there wouldn't be a 2 in front of $\delta^i{}_j$ in equation (2.14). The reason for the 2 factor is the 1/2 in the definition of the fiducial vector $\boldsymbol{p}$. The 1/2 in $\boldsymbol{p}$ in turn carries from the density operator formalism, from which the fiducial formalism derives. There we have that $\boldsymbol{\rho} = \frac{1}{2}\left(\mathbf{1} + r^x\boldsymbol{\sigma}_x + r^y\boldsymbol{\sigma}_y + r^z\boldsymbol{\sigma}_z\right)$ where $\boldsymbol{\sigma}_i$, $i = x,y,z$ are the Pauli matrices and $\mathbf{1}$ is the identity matrix.

$$\begin{aligned}
\boldsymbol{\omega}^1 &= \frac{1}{2}\left(\boldsymbol{\varsigma}^1 - \boldsymbol{\varsigma}^z\right) \\
\boldsymbol{\omega}^2 &= \frac{1}{2}\left(\boldsymbol{\varsigma}^1 + \boldsymbol{\varsigma}^x\right) \\
\boldsymbol{\omega}^3 &= \frac{1}{2}\left(\boldsymbol{\varsigma}^1 + \boldsymbol{\varsigma}^y\right)
\end{aligned} \tag{2.15}$$

and

$$\begin{aligned}
\boldsymbol{\varsigma}^1 &= \boldsymbol{\omega}^0 + \boldsymbol{\omega}^1 \\
\boldsymbol{\varsigma}^x &= -\boldsymbol{\omega}^0 - \boldsymbol{\omega}^1 + 2\boldsymbol{\omega}^2 \\
\boldsymbol{\varsigma}^y &= -\boldsymbol{\omega}^0 - \boldsymbol{\omega}^1 + 2\boldsymbol{\omega}^3 \\
\boldsymbol{\varsigma}^z &= \boldsymbol{\omega}^0 - \boldsymbol{\omega}^1
\end{aligned} \tag{2.16}$$

The bases of Pauli vectors and forms

Constructing a metric tensor in the fiducial space

But we can do more with Pauli forms and vectors, since they form natural bases in the fiducial vector and form spaces of the qubit.

The first thing we are going to do with Pauli forms and Pauli vectors is to form metric tensors in the fiducial vector and form spaces. The metric tensor $\boldsymbol{g}$ in the fiducial vector space is defined as follows:

$$\boldsymbol{g} = \frac{1}{2}\left(\boldsymbol{\varsigma}_1 \otimes \boldsymbol{\varsigma}_1 + \boldsymbol{\varsigma}_x \otimes \boldsymbol{\varsigma}_x + \boldsymbol{\varsigma}_y \otimes \boldsymbol{\varsigma}_y + \boldsymbol{\varsigma}_z \otimes \boldsymbol{\varsigma}_z\right) \tag{2.17}$$

Its counterpart in the fiducial form space, $\tilde{\boldsymbol{g}}$, is defined by

$$\tilde{\boldsymbol{g}} = \frac{1}{2}\left(\boldsymbol{\varsigma}^1 \otimes \boldsymbol{\varsigma}^1 + \boldsymbol{\varsigma}^x \otimes \boldsymbol{\varsigma}^x + \boldsymbol{\varsigma}^y \otimes \boldsymbol{\varsigma}^y + \boldsymbol{\varsigma}^z \otimes \boldsymbol{\varsigma}^z\right) \tag{2.18}$$

Metric tensors $\boldsymbol{g}$ and $\tilde{\boldsymbol{g}}$ can be used to convert vectors to forms and vice versa. Consider the following:

$$\begin{aligned}
\langle \tilde{\boldsymbol{g}}, \boldsymbol{\varsigma}_j \rangle &= \langle \frac{1}{2}\sum_{i=1,x,y,z} \boldsymbol{\varsigma}^i \otimes \boldsymbol{\varsigma}^i, \boldsymbol{\varsigma}_j \rangle \\
&= \frac{1}{2}\sum_{i=1,x,y,z} \boldsymbol{\varsigma}^i 2\delta^i{}_j = \boldsymbol{\varsigma}^j
\end{aligned}$$

Similarly

$$\begin{aligned}
\langle \boldsymbol{\varsigma}^i, \boldsymbol{g} \rangle &= \langle \boldsymbol{\varsigma}^i, \frac{1}{2}\sum_{j=1,x,y,z} \boldsymbol{\varsigma}_j \otimes \boldsymbol{\varsigma}_j \rangle \\
&= \frac{1}{2}\sum_{j=1,x,y,z} 2\delta^i{}_j \boldsymbol{\varsigma}_j = \boldsymbol{\varsigma}_i
\end{aligned}$$

Converting vectors to forms Without much ado we can use the above to convert the fiducial vector $\boldsymbol{p}$ to its dual fiducial form $\tilde{\boldsymbol{p}}$ [12]

Fiducial form of a qubit

$$
\begin{aligned}
\tilde{\boldsymbol{p}} &= \langle \tilde{\boldsymbol{g}}, \boldsymbol{p} \rangle \\
&= \frac{1}{2}\left(\boldsymbol{\varsigma}^1 + r^x \boldsymbol{\varsigma}^x + r^y \boldsymbol{\varsigma}^y + r^z \boldsymbol{\varsigma}^z\right) \\
&= \frac{1}{2}\left(1 - r^x - r^y + r^z, 1 - r^x - r^y - r^z, 2r^x, 2r^y\right) \quad (2.19)
\end{aligned}
$$

And it works in the other direction too

$$
\boldsymbol{p} = \langle \tilde{\boldsymbol{p}}, \boldsymbol{g} \rangle \quad (2.20)
$$

Metric tensor $\tilde{\boldsymbol{g}}$ can be used to evaluate length of vectors in the fiducial space, hence its name, "the metric". To evaluate the length of $\boldsymbol{p}$ proceed as follows:

Using $\tilde{\boldsymbol{g}}$ to measure the length of $\boldsymbol{p}$.

$$
\begin{aligned}
\langle \tilde{\boldsymbol{g}}, \boldsymbol{p} \otimes \boldsymbol{p} \rangle &= \langle \frac{1}{2} \sum_{i=1,x,y,z} \boldsymbol{\varsigma}^i \otimes \boldsymbol{\varsigma}^i, \boldsymbol{p} \otimes \boldsymbol{p} \rangle \quad (2.21) \\
&= \frac{1}{2}\left(1 \cdot 1 + r^x r^x + r^y r^y + r^z r^z\right) \quad (2.22) \\
&= \frac{1}{2}\left(1 + \boldsymbol{r} \cdot \boldsymbol{r}\right) \quad (2.23)
\end{aligned}
$$

The same can be obtained by evaluating $\langle \tilde{\boldsymbol{p}}, \boldsymbol{p} \rangle$ or

$$
\frac{1}{2}\left(1 - r^x - r^y + r^z, 1 - r^x - r^y - r^z, 2r^x, 2r^y\right) \cdot \frac{1}{2}\begin{pmatrix} 1 + r^z \\ 1 - r^z \\ 1 + r^x \\ 1 + r^y \end{pmatrix}
$$

For fully polarized states, the pure states, we have that $\boldsymbol{r} \cdot \boldsymbol{r} = 1$ and so

$\langle \tilde{\boldsymbol{p}}, \boldsymbol{p} \rangle = 1$ for fully polarized states

$$
\langle \tilde{\boldsymbol{p}}, \boldsymbol{p} \rangle = 1
$$

For completely chaotic states, i.e., states for which $\boldsymbol{r} = \boldsymbol{0}$

$\langle \tilde{\boldsymbol{p}}, \boldsymbol{p} \rangle = 1/2$ for fully chaotic states

$$
\langle \tilde{\boldsymbol{p}}, \boldsymbol{p} \rangle = \frac{1}{2}
$$

---

[12] The notation $r^x \boldsymbol{\varsigma}^x + r^y \boldsymbol{\varsigma}^y + r^z \boldsymbol{\varsigma}^z$ does not look quite as elegant as $r^x \boldsymbol{\varsigma}_x + r^y \boldsymbol{\varsigma}_y + r^z \boldsymbol{\varsigma}_z$ because indexes x, y, and z are on the same level, instead of being placed on the alternate levels, as we have emphasized in section 1.7. The reason for this is that form $\tilde{\boldsymbol{p}}$ has been generated by conversion from vector $\boldsymbol{p}$. Still, we can rescue the situation by lowering indexes on $\boldsymbol{r}$, i.e., by rewriting the $\boldsymbol{r}$ dependent part of the form as $r_x \boldsymbol{\varsigma}^x + r_y \boldsymbol{\varsigma}^y + r_z \boldsymbol{\varsigma}^z$. We can do this because in orthonormal (Cartesian) coordinates $r^i = r_i, i = x, y, z$.

Tensors $\boldsymbol{g}$ and $\tilde{\boldsymbol{g}}$ have matrix representations, which can be derived from fiducial representations of Pauli vectors and forms. The procedure is somewhat laborious, but what comes out eventually are two symmetric matrices:

$$\boldsymbol{g} \equiv \frac{1}{2} \begin{pmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix} \tag{2.24}$$

and

$$\tilde{\boldsymbol{g}} \equiv \frac{1}{2} \begin{pmatrix} 4 & 2 & -2 & -2 \\ 2 & 4 & -2 & -2 \\ -2 & -2 & 4 & 0 \\ -2 & -2 & 0 & 4 \end{pmatrix} \tag{2.25}$$

It can be checked easily that $\|\boldsymbol{g}\| = \|\tilde{\boldsymbol{g}}\|^{-1}$, i.e., that matrix that corresponds to $\boldsymbol{g}$ is the inverse of the matrix that corresponds to $\tilde{\boldsymbol{g}}$.

But there is a better way to see that $\boldsymbol{g}$ and $\tilde{\boldsymbol{g}}$ are each other's inverses. Consider $\langle \tilde{\boldsymbol{g}}, \boldsymbol{g} \rangle$ not contracted fully, but instead contracted on one vector and one form only. The results is

$$\boldsymbol{\delta} = \frac{1}{2} \left(\boldsymbol{\varsigma}^1 \otimes \boldsymbol{\varsigma}_1 + \boldsymbol{\varsigma}^x \otimes \boldsymbol{\varsigma}_x + \boldsymbol{\varsigma}^y \otimes \boldsymbol{\varsigma}_y + \boldsymbol{\varsigma}^z \otimes \boldsymbol{\varsigma}_z \right)$$

The reason there is only one $\frac{1}{2}$ in front is because the other one cancels with the factor of 2 thrown out by the contractions. It is now easy to see that this new object is simply the Kronecker delta or, in other words, the identity, because it converts an arbitrary fiducial vector $\boldsymbol{p}$ into itself. Indeed

$$\langle \boldsymbol{\delta}, \boldsymbol{p} \rangle = \frac{1}{2} \sum_{i=1,x,y,z} \langle \boldsymbol{\varsigma}^i, \boldsymbol{p} \rangle \boldsymbol{\varsigma}_i = \frac{1}{2} \sum_{i=1,x,y,z} 2 p^i \boldsymbol{\varsigma}_i = \boldsymbol{p}$$

Pauli vectors can be used to extract $\boldsymbol{r}$ from $\tilde{\boldsymbol{p}}$.   As we used Pauli forms $\boldsymbol{\varsigma}^1$ through $\boldsymbol{\varsigma}^z$ to extract vector $\boldsymbol{r}$ from $\boldsymbol{p}$, we can use Pauli vectors $\boldsymbol{\varsigma}_1$ through $\boldsymbol{\varsigma}_z$ to extract $\boldsymbol{r}$ from $\tilde{\boldsymbol{p}}$:

$$\langle \tilde{\boldsymbol{p}}, \boldsymbol{\varsigma}_1 \rangle = 1 \tag{2.26}$$

$$\langle \tilde{\boldsymbol{p}}, \boldsymbol{\varsigma}_x \rangle = r_x \tag{2.27}$$

$$\langle \tilde{\boldsymbol{p}}, \boldsymbol{\varsigma}_y \rangle = r_y \tag{2.28}$$

$$\langle \tilde{\boldsymbol{p}}, \boldsymbol{\varsigma}_z \rangle = r_z \tag{2.29}$$

2.7   The Hamiltonian form

Hamiltonian form defined   The Hamiltonian form for a qubit is given by:[13] [14]

$$\boldsymbol{\eta} = -\mu \left(B_x \boldsymbol{\varsigma}^x + B_y \boldsymbol{\varsigma}^y + B_z \boldsymbol{\varsigma}^z \right) \tag{2.30}$$

If the qubit is implemented as a beam of neutrons, then μ is the magnetic moment of the neutron and $\boldsymbol{B} = B_x \boldsymbol{e}_x + B_y \boldsymbol{e}_y + B_z \boldsymbol{e}_z$ is the magnetic field. But a qubit may be implemented in other ways too, and then the physical meaning of μ and $\boldsymbol{B}$ is bound to be different. Yet the Hamiltonian form still looks the same.

Coefficients B_x, B_y and B_z can be extracted from the Hamiltonian form $\boldsymbol{\eta}$ by the application of Pauli vectors to it:

Extracting $\boldsymbol{B}$ from $\boldsymbol{\eta}$ with Pauli vectors

$$
\begin{aligned}
\langle \boldsymbol{\eta}, \boldsymbol{\varsigma}_1 \rangle &= 0 \\
\langle \boldsymbol{\eta}, \boldsymbol{\varsigma}_x \rangle &= -2\mu B_x \\
\langle \boldsymbol{\eta}, \boldsymbol{\varsigma}_y \rangle &= -2\mu B_y \\
\langle \boldsymbol{\eta}, \boldsymbol{\varsigma}_z \rangle &= -2\mu B_z
\end{aligned}
$$

The reason we get $-2\mu B_x$ instead of just $-\mu B_x$ is because there is no $1/2$ in front of $\boldsymbol{\eta}$, and remember that $\langle \boldsymbol{\varsigma}^i, \boldsymbol{\varsigma}_i \rangle = 2$ for $i = x, y, z$.

The Hamiltonian form is used to calculate the average value of the energy over the statistical ensemble of qubits described by the fiducial vector $\boldsymbol{p}$. The formula[15] looks the same as formulas we have derived for classical randomly fluctuating registers in section 1.6, page 22:

The Hamiltonian form is used to calculate the average energy over the statistical ensemble of qubits.

$$
\begin{aligned}
\langle E \rangle &= \langle \boldsymbol{\eta}, \boldsymbol{p} \rangle \\
&= \left\langle -\mu \left(B_x \boldsymbol{\varsigma}^x + B_y \boldsymbol{\varsigma}^y + B_z \boldsymbol{\varsigma}^z \right), \frac{1}{2} \left(\boldsymbol{\varsigma}_1 + r^x \boldsymbol{\varsigma}_x + r^y \boldsymbol{\varsigma}_y + r^z \boldsymbol{\varsigma}_z \right) \right\rangle \\
&= -\mu \left(B_x r^x + B_y r^y + B_z r^z \right)
\end{aligned}
$$

This expression looks exactly like the classical expression that describes the energy of magnetic dipole $\boldsymbol{\mu}$ in magnetic field $\boldsymbol{B}$

The quantum formula is strikingly similar to the classical formula that describes the energy of a magnetic dipole in a magnetic field.

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B} \tag{2.31}$$

---

[13]Here we keep indexes x, y and z on $\boldsymbol{B}$ down, because we really want to use $\boldsymbol{B}$ as a form, not as a vector. This makes no difference computationally, because $B_i = B^i, i = x, y, z$ in orthonormal (Cartesian) coordinates. See footnote 12 on page 65.

[14] There is a reason why we have chosen $\boldsymbol{\eta}$ for the Hamiltonian form. The usual letter used for the Hamiltonian operator in quantum mechanics is $\boldsymbol{H}$. But we need a low case Greek letter here. Since H happens to be the Greek capital letter version of "eta", the low case of which is η, we end up with $\boldsymbol{\eta}$ for the Hamiltonian *form*. The Greek word from which the word "energy" derives is $\varepsilon\nu\varepsilon\rho\gamma\varepsilon\iota\alpha$, which means "activity" or "effect". It begins with ε, not with η.

[15]This formula is equivalent to $\langle E \rangle = \text{Tr}\,(\boldsymbol{H}\boldsymbol{\rho})$, where $\boldsymbol{H}$ is the Hamiltonian operator and $\boldsymbol{\rho}$ is the density operator.

if we identify

$$\boldsymbol{\mu} = -\mu \begin{pmatrix} r^x \\ r^y \\ r^z \end{pmatrix} \tag{2.32}$$

But we must remember that in the world of quantum mechanics $\boldsymbol{r}$ is the vector that parameterizes probability measurements made on the statistical ensemble of, e.g., neutrons, and does not represent the space orientation of the individual neutron. The give-aways that separate the quantum formula from its classical cousin are the angular brackets around the E.

Hamiltonian As we could convert the fiducial vector $\boldsymbol{p}$ to the fiducial form $\tilde{\boldsymbol{p}}$ by contracting $\boldsymbol{p}$
vector with metric $\tilde{\boldsymbol{g}}$, similarly we can convert the Hamiltonian form to the Hamiltonian vector by contracting the form with metric $\boldsymbol{g}$:

$$\tilde{\boldsymbol{\eta}} = \langle \boldsymbol{\eta}, \boldsymbol{g} \rangle = -\mu \left(B_x \boldsymbol{\varsigma}_x + B_y \boldsymbol{\varsigma}_y + B_z \boldsymbol{\varsigma}_z \right) \tag{2.33}$$

The measurement of average energy on the statistical ensemble that corresponds to $\boldsymbol{p}$ can be then expressed also in this way:

$$\langle E \rangle = \langle \tilde{\boldsymbol{p}}, \tilde{\boldsymbol{\eta}} \rangle \tag{2.34}$$

Interchangeability Forms and vectors are interchangeable – as long as we remember to interchange
of forms and both at the same time!
vectors   It is instructive to evaluate the Hamiltonian form on the basis states $|\uparrow\rangle$ and
Energy of the $|\downarrow\rangle$. Remember that neutrons in the beam always align or counter-align with the
basis states direction of the magnetic field. The basis states $|\uparrow\rangle$ and $|\downarrow\rangle$ can therefore be observed if a neutron is placed in $\boldsymbol{B} = B\boldsymbol{e}_z$. For the $|\uparrow\rangle$ state we have $\boldsymbol{r} = \boldsymbol{e}_z$ and for the $|\downarrow\rangle$ state we have $\boldsymbol{r} = -\boldsymbol{e}_z$. Hence

$$\langle \boldsymbol{\eta}, \uparrow \rangle = -\mu B \tag{2.35}$$

and

$$\langle \boldsymbol{\eta}, \downarrow \rangle = \mu B \tag{2.36}$$

The energy difference between the two states is

$$\Delta E = E_{|\downarrow\rangle} - E_{|\uparrow\rangle} = 2\mu B \tag{2.37}$$

Suppose we take a neutron in the $|\downarrow\rangle$ state and stick it in the chamber filled with $\boldsymbol{B} = B\boldsymbol{e}_z$. Initially the neutron is going to maintain its original orientation of the magnetic moment. But this orientation is not stable, because there is another orientation possible with lower energy, namely

$|\uparrow\rangle$. Consequently the neutron will eventually flip to the lower energy state, releasing a photon of energy $2\mu B$ in the process.

The photon release and flip is called *dissipation*. It is quite unpredictable al- *Energy* though the so called *half-life* of the higher energy state can be measured for various *dissipation* circumstances. Dissipation is amongst the main causes of errors in quantum computing, so a better understanding of this phenomenon is clearly called for. We will present a simplified model of dissipation in chapter 5.

The frequency of light emitted in the dissipation process is given by

$$\Delta E = 2\mu B = \hbar\omega$$

which yields

$$\omega = \frac{2\mu B}{\hbar} \tag{2.38}$$

The important thing to note here is that we never observe neutrons emitting smaller portions of energy in this process (which would result in $\omega < 2\mu B/\hbar$). This means that the flip does not proceed by going through some intermediate states, which we could interpret as various angles of tilt between "up" and "down". The neutron switches instantaneously between the two opposite and discrete configurations *without* going through intermediate states, because ... there aren't any to go through. If there were, we would have seen them in the beam splitting experiment.

2.8   Qubit evolution

Amongst most astonishing discoveries of quantum mechanics is that whereas the *Quantum* behavior of any individual microscopic system is chaotic and in general unpre- *probability* dictable, yet probability distributions that pertain to microscopic systems can be *distributions* described with great precision and their evolution predicted accurately by the means *evolve* of deterministic differential equations. *deterministicly*

> This is not an entirely new situation in physics. For example the classical diffusion equation is deterministic yet it can be derived from a microscopic picture that assumes completely random molecular motion.
> Yet in quantum mechanics the question "where do quantum probabilities come from" remains unanswered[16] and while the orthodox pronouncement

---

[16]Although the question "where quantum probabilities come from" remains unanswered, there have been various answers proposed. These range from the de Broglie-Bohm pilot wave theory [8] through the Everett's many-worlds interpretation of quantum mechanics [13] and more. The reason for our statement that this remains an unsolved problem is that none of the proposed solutions, which are all mathematically sound, have been verified experimentally.

states that quantum probabilities are fundamental and unexplainable, most physicists and chemists who work with quantum mechanics daily have a vague picture in their mind that associates the wave function of a particle with a sort of tension traveling through space – wherever the tension is greater the particle is more likely to materialize. In the absence of macroscopic environments the particle is dissolved in the "tension wave" and non-local, but as soon as it encounters a "measuring apparatus" (it can as well be a large lump of matter without any dials, the particle doesn't care) the whole shebang shrinks and, pronto, we get a point-like energy transfer: the particle has been registered.

Professional physicists and chemists are usually very careful not to divulge, especially to their colleagues and students, their feelings on this matter [4] and just stick to the lore. It is only in the most intimate moments of marital bliss that a spouse may overhear the physicist or the chemist uttering an illicit thought about quantum ontology in their sleep or under shower.

Limited applicability of the Schrödinger equation

The equation we are going to introduce in this section does not describe every aspect of microscopic behavior. For example, it does not describe the measurement process or the dissipation process. In order to analyze these processes we will have to advance to multi-qubit systems, because both derive from interactions of a qubit with its environment. The environment, that can be modeled by other qubits our selected qubit interacts with. In simple though quite informative models it is enough to add just one more qubit to the description.

The evolution of a single qubit in presence of a possibly varying "magnetic field" $\boldsymbol{B}$ is given by the following equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p} = \frac{1}{2\hbar}\langle \boldsymbol{\eta} \otimes \tilde{\boldsymbol{p}}, \boldsymbol{\epsilon}\rangle \tag{2.39}$$

where

$$\boldsymbol{\epsilon} = \sum_{i,j,k\in\{x,y,z\}} \epsilon^{ijk}\boldsymbol{\varsigma}_i \otimes \boldsymbol{\varsigma}_j \otimes \boldsymbol{\varsigma}_k \tag{2.40}$$

where $\epsilon^{ijk} = +1$ for $i = x$, $j = y$, $k = z$ and even permutations thereof, -1 for odd permutations and 0 otherwise.

The pedigree of equation (2.39)

Equation (2.39) is due to many people, beginning with Schrödinger who *guessed* its form for the *wave function* (already mentioned in the aside above). The Schrödinger equation was later enhanced by Pauli to describe wave functions of magnetized particles. Then von Neumann modified the Schrödinger equation still further to describe evolution of the *density operator* (mentioned in previous sections). Finally, equation (2.39) was obtained from the von Neumann equation by taking special projections that yielded directly measurable probabilities p^0 through p^3 [26] [53].

At the end of the day, all these equations describe exactly the same phenomena, but equation (2.39) yields results that can be interpreted directly and without a lot of additional gymnastics. Because the original Schrödinger equation as well as its Pauli extension were simply guessed and cannot be *derived* from simpler assumptions by a meaningful reasoning[17] and because equation (2.39) is equivalent to the Schrödinger-Pauli-von-Neumann equation for a qubit, we are basically going to stop right here. We can just as well say that equation (2.39) has been guessed and defer its verification to laboratory experiments.

Tensor $\boldsymbol{\epsilon}$ given by equation (2.40) was invented by Levi-Civita. It is the same kind *Levi-Civita* of a device that is used in calculating the *vector* or *cross* product of two vectors. *tensor* We will see shortly that equation (2.39) translates easily into equation for $\mathrm{d}\boldsymbol{r}/\mathrm{d}t$ in which the cross product appears on the right hand side.

Let us substitute our generic parameterizations for $\boldsymbol{p}$ and $\boldsymbol{\eta}$ in equation (2.39). Recall that

$$\boldsymbol{p} = \frac{1}{2}\left(\boldsymbol{\varsigma}_1 + r^x\boldsymbol{\varsigma}_x + r^y\boldsymbol{\varsigma}_y + r^z\boldsymbol{\varsigma}_z\right)$$

Substituting this into the left hand side of equation (2.39) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p} = \frac{1}{2}\frac{\mathrm{d}r^x}{\mathrm{d}t}\,\boldsymbol{\varsigma}_x + \frac{1}{2}\frac{\mathrm{d}r^y}{\mathrm{d}t}\,\boldsymbol{\varsigma}_y + \frac{1}{2}\frac{\mathrm{d}r^z}{\mathrm{d}t}\,\boldsymbol{\varsigma}_z \qquad (2.41)$$

The term proportional to $\boldsymbol{\varsigma}_1$ is constant and so its time derivative vanishes.

The devil is in the right hand side of equation (2.39), but we can outwit him by recalling that

$$\langle \boldsymbol{\eta}, \boldsymbol{\varsigma}_i \rangle = -2\mu B_i \quad \text{for} \quad i = x, y, z$$

and

$$\langle \tilde{\boldsymbol{p}}, \boldsymbol{\varsigma}_i \rangle = r_i \quad \text{for} \quad i = x, y, z$$

Plugging these formulae into the right hand side of equation (2.39) and using the notational shortcut $\boldsymbol{\varsigma}_x \wedge \boldsymbol{\varsigma}_y = \boldsymbol{\varsigma}_x \otimes \boldsymbol{\varsigma}_y - \boldsymbol{\varsigma}_y \otimes \boldsymbol{\varsigma}_x$ yields

$$\begin{aligned}
&\frac{1}{2\hbar}\langle \boldsymbol{\eta} \otimes \tilde{\boldsymbol{p}}, \boldsymbol{\epsilon}\rangle \\
&= \frac{1}{2\hbar}\left\langle \boldsymbol{\eta} \otimes \tilde{\boldsymbol{p}}, (\boldsymbol{\varsigma}_x \wedge \boldsymbol{\varsigma}_y) \otimes \boldsymbol{\varsigma}_z + (\boldsymbol{\varsigma}_y \wedge \boldsymbol{\varsigma}_z) \otimes \boldsymbol{\varsigma}_x + (\boldsymbol{\varsigma}_z \wedge \boldsymbol{\varsigma}_x) \otimes \boldsymbol{\varsigma}_y\right\rangle \\
&= -\frac{\mu}{\hbar}\left((B_x r_y - B_y r_x)\,\boldsymbol{\varsigma}_z + (B_y r_z - B_z r_y)\,\boldsymbol{\varsigma}_x + (B_z r_x - B_x r_z)\,\boldsymbol{\varsigma}_y\right) \\
&= \frac{\mu}{\hbar}\left((\boldsymbol{r} \times \boldsymbol{B})^x\,\boldsymbol{\varsigma}_x + (\boldsymbol{r} \times \boldsymbol{B})^y\,\boldsymbol{\varsigma}_y + (\boldsymbol{r} \times \boldsymbol{B})^z\,\boldsymbol{\varsigma}_z\right) \qquad (2.42)
\end{aligned}$$

---

[17]The Schrödinger equation can be derived from Feynman rules about path integrals but it is still unclear after 55 years whether the reasoning in this case is meaningful, even though Feynman path integrals are used routinely to construct various quantum field theories and to carry out numerical computations on discrete lattices.

Finally, combining equations (2.41) and (2.42) results in

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r} = \frac{2\mu}{\hbar}\boldsymbol{r} \times \boldsymbol{B} \qquad (2.43)$$

Equation (2.43) is basically the same as equation (2.39), but it is easier to read, solve and interpret and it doesn't use funny symbols such as $\langle, \rangle$, $\tilde{\boldsymbol{p}}$ and $\boldsymbol{\epsilon}$.

Equation (2.43) preserves the length of $\boldsymbol{r}$.

The first thing to notice about it is that it does not change the length of $\boldsymbol{r}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{r} \cdot \boldsymbol{r}) = 2\boldsymbol{r} \cdot \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r} = \frac{4\mu}{\hbar}\boldsymbol{r} \cdot (\boldsymbol{r} \times \boldsymbol{B}) = 0 \qquad (2.44)$$

on account of $\boldsymbol{r} \times \boldsymbol{B}$ being perpendicular both to $\boldsymbol{r}$ and $\boldsymbol{B}$. This means that solutions to equation (2.43) will be given as rotations of vector $\boldsymbol{r}$. If we begin with a pure, i.e., fully polarized state, for which $\boldsymbol{r} \cdot \boldsymbol{r} = 1$, the state will remain fully polarized. This is why we can say outright that equation (2.43) cannot describe measurements, because these, in general, reduce the length of $\boldsymbol{r}$.

Solutions of (2.43) define quantum computational operations.

In the next two sections we will see that these rotations can be quite complex, but at the same time, we will encounter a great deal of very interesting physics, which is directly applicable to quantum computing.

2.9 Larmor precession

Consider a qubit described by the fiducial vector $\boldsymbol{p}$ in presence of the static (i.e., unchanging in time) and uniform (i.e., unchanging in space) field $\boldsymbol{B} = B\boldsymbol{e}_z$. Substituting $B\boldsymbol{e}_z$ in place of $\boldsymbol{B}$ in equation (2.43) results in

$$\frac{\mathrm{d}r^x}{\mathrm{d}t} = \frac{2\mu}{\hbar}r^y B \qquad (2.45)$$

$$\frac{\mathrm{d}r^y}{\mathrm{d}t} = -\frac{2\mu}{\hbar}r^x B \qquad (2.46)$$

$$\frac{\mathrm{d}r^z}{\mathrm{d}t} = 0 \qquad (2.47)$$

We will assume that the initial state of the qubit is fully polarized and so, as equation (2.44) tells us, the qubit will remain fully polarized.

Larmor precession does not affect p^0 and p^1. Equations (2.45) and (2.46) are analogous to equations that describe the harmonic oscillator.

Equation (2.47) is easiest to solve. It says that $r^z = \cos\theta$ is constant (hence θ is constant too), which implies that the probabilities of finding the qubit in states $|\uparrow\rangle$ and $|\downarrow\rangle$ are constant too.

Equations (2.45) and (2.46) are coupled in the same way that the position and

the velocity are coupled in the equations of the harmonic oscillator. To see this let us take the second derivative of r^x with respect to time.

$$\frac{\mathrm{d}^2 r^x}{\mathrm{d}t^2} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}r^x}{\mathrm{d}t} = \frac{2\mu B}{\hbar}\frac{\mathrm{d}r^y}{\mathrm{d}t} = -\left(\frac{2\mu B}{\hbar}\right)^2 r^x \tag{2.48}$$

Similarly for r^y

$$\frac{\mathrm{d}^2 r^y}{\mathrm{d}t^2} = -\left(\frac{2\mu B}{\hbar}\right)^2 r^y$$

Equation (2.48) can be satisfied by the following ansatz

$$r^x(t) = r_0^x \cos\omega_L t$$

Larmor frequency

which assumes implicitly that at $t = 0$ $r^x = r_0^x = \sin\theta$ and $\mathrm{d}r^x/\mathrm{d}t = 0$. The constant ω_L is given by

$$\omega_L = \frac{2\mu B}{\hbar} \tag{2.49}$$

and is called the *Larmor frequency* after Irish physicist, Sir Joseph Larmor (1857-1942), who was the first to explain the splitting of spectral lines by a magnetic field.

r^y can be obtained from equation (2.45), which states that

$$r^y = \frac{1}{\omega_L}\frac{\mathrm{d}}{\mathrm{d}t}r^x = -\frac{1}{\omega_L}\omega_L r_0^x \sin\omega_L t = -r_0^x \sin\omega_L t$$

In summary, this is our solution

$$r^x = \sin\theta\cos\omega_L t \tag{2.50}$$
$$r^y = -\sin\theta\sin\omega_L t \tag{2.51}$$
$$r^z = \cos\theta \tag{2.52}$$

The solution describes vector $\boldsymbol{r}$, which rotates about the direction of $\boldsymbol{B}$ in such a way that its projection on the direction of $\boldsymbol{B}$ stays the same, see figure 2.6. Such a motion is called a precession, hence this phenomenon is referred to as the Larmor precession.

Larmor precession

Although the probabilities of finding the qubit in $|\uparrow\rangle$ and $|\downarrow\rangle$ are constant, the probabilities of finding the qubit in states $|\rightarrow\rangle$ and $|\otimes\rangle$ change all the time. The state remains fully polarized, i.e., pure, but the direction of its polarization precesses around $\boldsymbol{B}$.
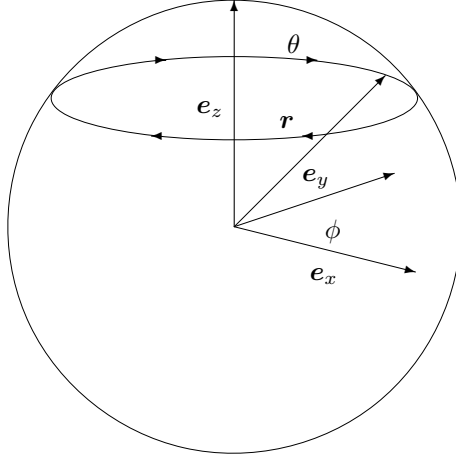
Figure 2.6: Larmor precession

The Larmor precession occurs with the angular frequency of $\omega_L = 2\mu B/\hbar$, which is the same as the frequency of the photon emitted as the result of the dissipative transition from $|\uparrow\rangle$ to $|\downarrow\rangle$ (cf. equation (2.38) on page 69).

In order to evaluate probabilities of finding the qubit in various basis states we *Finding the* have to substitute solutions (2.50), (2.51) and (2.52) into the fiducial vector $\boldsymbol{p}$.   *probabilities*

$1 + \cos\alpha =$ The probability of finding the qubit in state $|\uparrow\rangle$ is
$2\cos^2\frac{\alpha}{2}$

$$p^0 = \frac{1}{2}\left(1 + r^z\right) = \frac{1}{2}\left(1 + \cos\theta\right) = \cos^2\frac{\theta}{2}$$

$1 - \cos\alpha =$ The probability of finding the qubit in state $|\downarrow\rangle$ is
$2\sin^2\frac{\alpha}{2}$

$$p^1 = \frac{1}{2}\left(1 - r^z\right) = \frac{1}{2}\left(1 - \cos\theta\right) = \sin^2\frac{\theta}{2}$$

The probability of finding the qubit in state $|\rightarrow\rangle$ is

$$p^2 = \frac{1}{2}\left(1 + r^x\right) = \frac{1}{2}\left(1 + \sin\theta\cos\omega_L t\right)$$

and the probability of finding it in state $|\otimes\rangle$ is

$$p^3 = \frac{1}{2}\left(1 + r^y\right) = \frac{1}{2}\left(1 - \sin\theta\sin\omega_L t\right)$$

Larmor   Note that just placing the qubit in a static and uniform magnetic field $\boldsymbol{B} = B\boldsymbol{e}_z$
precession does
not measure the
qubit.

does not in itself result in the measurement, which is why the qubit remains in a pure state, although the state itself varies with time. In order to measure the qubit we have to pass it through a chamber where magnetic field has a non-vanishing gradient. Another way to measure the qubit is to measure its own magnetic field with very sensitive Helmholtz coils, which is how the results of computations are read out in Nuclear Magnetic Resonance experiments.

If the qubit is implemented by other means, i.e., not as a microscopic magnet, the physical meaning of $\boldsymbol{B}$ and μ are different and then appropriate measurement methods have to be devised.

The Larmor precession can be a parasitic effect in context of quantum computing. *Larmor* Usually we want our qubits to just stay as they are. It is sometimes possible to just *precession as a* switch $\boldsymbol{B}$ off between applications of various quantum gates. This, again, depends *parasitic effect* on how qubits are implemented. In Nuclear Magnetic Resonance experiments $\boldsymbol{B}$ cannot be switched off. In this case an elaborate procedure called refocusing must *Refocusing* be deployed to cancel the effects of Larmor precession in final account.

2.10   Rabi oscillations

The Larmor precession does not change the proportions of $|\uparrow\rangle$ to $|\downarrow\rangle$ in a qubit state *Larmor* described by $\boldsymbol{p}$. If we were to make the associations $|\uparrow\rangle \equiv |\,0\rangle$ and $|\downarrow\rangle \equiv |\,1\rangle$ then the *precession does* precessing qubit would stay put, as far as its computational value is concerned, even *not affect the* though something would keep going inside it, so to speak. But Larmor precession *computational* may have effect on computations in case the qubit couples to other qubits, so it *value of a qubit.* cannot be ignored.

The question for this section is how we can change the proportion of $|\uparrow\rangle$ to $|\downarrow\rangle$ in *How to flip a* a qubit, and, in particular, how we can flip $|\uparrow\rangle$ into $|\downarrow\rangle$ and vice versa in a controlled *qubit?* way, i.e., without waiting for a dissipation event to occur.

There are various ways in which this can be accomplished. The simplest and the most commonly practiced is to apply a small magnetic oscillation in the plane perpendicular to the background magnetic field $\boldsymbol{B}$. The result of this oscillation will be a slow, very slow in fact compared with the Larmor precession, latitudinal movement of the tip of vector $\boldsymbol{r}$. The latitudinal drift when combined with the Larmor precession results in drawing a spiral curve on the surface of the Bloch sphere that connects its two poles.

In order to analyze this effect in more detail we are going to solve equation (2.43) yet again, taking all components of $\boldsymbol{B}$ into account, although we are still going to

make some simplifying assumptions. Expanding $\mathrm{d}\boldsymbol{r}/\mathrm{d}t = (2\mu/\hbar)\,\boldsymbol{r} \times \boldsymbol{B}$ in the x, y, z coordinates explicitly gives us the following three equations

$$\frac{\mathrm{d}r^x}{\mathrm{d}t} = \frac{2\mu}{\hbar}\left(r^y B^z - r^z B^y\right) \tag{2.53}$$

$$\frac{\mathrm{d}r^y}{\mathrm{d}t} = \frac{2\mu}{\hbar}\left(r^z B^x - r^x B^z\right) \tag{2.54}$$

$$\frac{\mathrm{d}r^z}{\mathrm{d}t} = \frac{2\mu}{\hbar}\left(r^x B^y - r^y B^x\right) \tag{2.55}$$

Approximations made   We are going to assume that

$$
\begin{aligned}
|B^x| &\ll |B^z| \\
|B^y| &\ll |B^z| \\
B^z &= \text{const}
\end{aligned}
$$

We are also going to assume that the initial state is fully polarized, which implies that the solution for all values of t must be fully polarized too.

Since B^x and B^y are small compared to B^z we can try the following form of the

General form of the solution   solution

$$r^x = \sin\theta(t)\cos\omega_L t \tag{2.56}$$

$$r^y = -\sin\theta(t)\sin\omega_L t \tag{2.57}$$

$$r^z = \cos\theta(t) \tag{2.58}$$

In other words, we assume that the qubit keeps precessing as in section 2.9, but this time the angle θ is no longer constant. Instead it is a slowly varying function of time – slowly compared to $\omega_L t$. This assumption is equivalent to saying that we are going to ignore $r^z B^y$ compared to $r^y B^z$ in equation (2.53) and $r^z B^x$ compared to $r^x B^z$ in equation (2.54). With these simplifications equations (2.53) and (2.54) are the same as equations (2.45) and (2.46) on page 72, which, as we know already, describe the Larmor precession. So this leaves us with equation (2.55).

LHS worked on   First let us substitute (2.58) in the left hand side of equation (2.55). The result is

$$\frac{\mathrm{d}}{\mathrm{d}t}r^z = \frac{\mathrm{d}}{\mathrm{d}t}\cos\theta(t) = -\sin\theta(t)\frac{\mathrm{d}}{\mathrm{d}t}\theta(t) \tag{2.59}$$

RHS worked on   Substituting (2.56) and (2.57) in the right hand side of equation (2.55) yields

$$\frac{2\mu}{\hbar}\left(r^x B^y - r^y B^x\right)$$

$$= \frac{2\mu}{\hbar}\left(B^y \sin\theta(t)\cos\omega_L t + B^x \sin\theta(t)\sin\omega_L t\right) \tag{2.60}$$

Comparing equations (2.59) and (2.60) tells us that we can cancel $\sin\theta(t)$ that occurs on both sides and this yields a simpler differential equation for $\theta(t)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta(t) = \frac{2\mu}{\hbar}\left(B^x\sin\omega_L t + B^y\cos\omega_L t\right) \tag{2.61}$$

Until now we have not specified B^x and B^y, other than to say that they're much *Specifications* smaller than B^z. We are now going to specify both, so as to make equation (2.61) *for B^x and B^y* easier to solve. And so, let us assume the following

$$B^x = B_\perp \sin\omega t \tag{2.62}$$
$$B^y = B_\perp \cos\omega t \tag{2.63}$$

On this occasion we shall also rename B^z to $B_\parallel$. Plugging (2.62) and (2.63) into (2.61) results in

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta(t) = \frac{2\mu B_\perp}{\hbar}\left(\sin\omega_L t\sin\omega t + \cos\omega_L t\cos\omega t\right)$$

Now we can invoke the well known high school trigonometric formula (shown in the margin note) to wrap this into

$\cos\alpha\cos\beta +$
$\sin\alpha\sin\beta =$

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta(t) = \frac{2\mu B_\perp}{\hbar}\cos\left(\omega_L - \omega\right)t \tag{2.64}$$

$\cos\left(\alpha - \beta\right)$

which at long last can be solved easily for $\theta(t)$ *Solution found*

$$\theta(t) = \frac{2\mu B_\perp}{\hbar}\frac{\sin\left(\omega_L - \omega\right)t}{\omega_L - \omega} \tag{2.65}$$

assuming that $\theta(0) = 0$, i.e., assuming that at $t = 0$ the qubit is in state $|\uparrow\rangle$.

2.10.1   Solution at resonance

We will first consider solutions at frequencies ω that are very close to ω_L or right *Resonance* on the spot, i.e, $\omega = \omega_L$. This corresponds to buzzing the qubit with a small *condition* transverse magnetic field $\boldsymbol{B}_\perp = \left(B_\perp\sin\omega_L t\right)\boldsymbol{e}_x + \left(B_\perp\cos\omega_L t\right)\boldsymbol{e}_y$ that rotates with the Larmor frequency of the qubit itself.

For very small values of $\left(\omega_L - \omega\right)$ we can replace the sine function with $\lim\limits_{x\to 0}\sin x = x$

$$\sin\left(\omega_L - \omega\right)t \approx \left(\omega_L - \omega\right)t$$

This way the $\left(\omega_L - \omega\right)$ factor cancels out and equation (2.65) becomes   *Solution at resonance*

$$\theta(t) = \frac{2\mu B_\perp}{\hbar} t \tag{2.66}$$

Angle θ grows linearly with time t. Although θ is restricted to $0 \leq \theta \leq \pi$ in   *Rabi oscillations*
principle, the solution doesn't care about this and continues to grow beyond π.
What does this mean? First, note that $\cos\theta = \cos(2\pi - \theta)$. This implies that as
far as r^z is concerned, once θ has exceeded π, we can replace it with $\theta' = 2\pi - \theta$
and reduce the angle to the $0 \ldots \pi$ range. In other words, once θ has reached π, it
begins to swing back, until it returns to $\theta = 0$, whereupon it is going to resume its
journey south.

So, what we have here are oscillations of vector r between north and south
interposed on Larmor precession. The north-south oscillations of r produced by this
process, see figure 2.7, are called Rabi oscillations, after U. S. physicist and Nobel
prize winner (1944) Isidor Isaac Rabi (1898-1988), and occur with the frequency
that is much lower than the Larmor frequency.
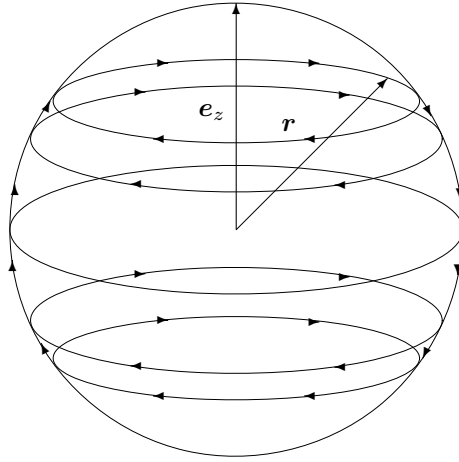
Figure 2.7: Rabi oscillations – as vector r continues to precess very rapidly about
e_z, its latitude θ (measured from the north pole) increases very slowly, until r flips
to the southern hemisphere and approaches the south pole, whereupon the process
reverses.

Rabi frequency   The period of these oscillations is given by

$$\theta(T_R) = 2\pi = \frac{2\mu B_\perp}{\hbar} T_R$$

which yields

$$T_R = \frac{\pi\hbar}{\mu B_\perp} \tag{2.67}$$

The angular frequency of Rabi oscillations, ω_R, is

$$\omega_R = \frac{2\pi}{T_R} = \frac{2\mu B_\perp}{\hbar} \ll \frac{2\mu B_\parallel}{\hbar} = \omega_L \tag{2.68}$$

because

$$|B_\perp| \ll |B_\parallel|$$

If we start the process with a qubit in the basis state $|\uparrow\rangle = |\,0\rangle$, aligned with *How long does it* $\boldsymbol{B}_\parallel$, it is going to take $\Delta t = T_R/2 = \pi\hbar/(2\mu B_\perp)$ of buzzing the qubit with $\boldsymbol{B}_\perp = $ *take to flip a* $(B_\perp \sin\omega t)\,\boldsymbol{e}_x + (B_\perp \cos\omega t)\,\boldsymbol{e}_y$ to flip it to $|\downarrow\rangle = |\,1\rangle$. And similarly, to flip the qubit *qubit?* from $|\downarrow\rangle = |\,1\rangle$ to $|\uparrow\rangle = |\,0\rangle$ will take the same amount of time.

The full form of the solution for $\boldsymbol{r}(t)$ is

$$r^x = \sin\omega_R t \cos\omega_L t \tag{2.69}$$
$$r^y = -\sin\omega_R t \sin\omega_L t \tag{2.70}$$
$$r^z = \cos\omega_R t \tag{2.71}$$

Let us have another close look at what is going to happen when θ grows *More about $\boldsymbol{r}$* above π. On the way back from the south pole $\pi < \theta < 2\pi$ and in this region *going back north* $\sin\theta$ is negative. Replacing θ with $\theta' = 2\pi - \theta$, which has the effect of reducing θ back to $[0, \pi]$, and writing the minus in front of $\sin\theta$ explicitly, we find that the solution assumes the following form now:

$$r^x = -\sin\theta' \cos\omega_L t = \sin\theta' \cos(\omega_L t + \pi)$$
$$r^y = \sin\theta' \sin\omega_L t = -\sin\theta' \sin(\omega_L t + \pi)$$
$$r^z = \cos\theta'$$

where we have made use of the fact that $-\sin\alpha = \sin(\alpha + \pi)$ and $-\cos\alpha = \cos(\alpha + \pi)$ in order to absorb the changed sign into the sin and cos of the Larmor precession.

This result tells us that on its way back north vector $\boldsymbol{r}$ is going to visit a point of the Bloch sphere, which, for a given angle θ, is on the other side, compared to the point it had visited for the same angle θ on the way south.

As vector $\boldsymbol{r}$ nears the north pole, its returning trajectory on the Bloch sphere is going to align itself with the starting trajectory, because it is its reflection with respect to the $\boldsymbol{e}_z$ axis, so that after having reached the north pole vector $\boldsymbol{r}$ will go south on exactly the same trajectory it traced originally.

As we go south and north, and south again and north again, we're going to follow the same spiral all the time, crossing the equator at exactly the

same two points, the point on the return voyage being on the other side, with respect to the point at which we cross the equator on the way south.

This property can be used in the following interesting operation, which is due to U. S. physicist and Nobel prize winner (1989) Norman Foster Ramsey (born in 1915).

Suppose we use the Rabi oscillations to tilt r by $90°$ south. We can now switch off the buzzing field $B_\perp$ and just wait a while, allowing for the Larmor precession to rotate r about the z axis. Then we can switch $B_\perp$ on again and continue tilting the qubit. If the Larmor precession has rotated the qubit by a multiple of 2π (i.e., $360°$) in the meantime, the qubit will come back to the point from which it left and on receiving the buzzing signal will resume its journey to the south pole. But if the Larmor precession has rotated the qubit by an odd multiple of π (i.e., $180°$), the qubit will resume its journey along the Rabi trajectory at the point that's on the other side of the Bloch sphere and so it'll come back north instead.

What is going to happen if the Larmor precession leaves vector r stranded at some point that is between the two points where the Rabi spiral crosses the equator? When the buzzing signal kicks in, it is going to be out of phase now with respect to the Larmor precession. We can describe this by altering equation (2.64) on page 77 and adding a fixed angle ϕ to the phase of the qubit:

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta(t) = \omega_R \cos\left((\omega_L - \omega)\, t + \phi\right)$$

Making again use of the school formula shown in the margin we can rewrite this as

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta(t) = \omega_R \left(\cos\left(\omega_L - \omega\right) t \cos\phi - \sin\left(\omega_L - \omega\right) t \sin\phi\right)$$

The resonance condition $\omega = \omega_L$ kills the $\sin(\omega_L - \omega)t$ term and converts $\cos(\omega_L - \omega)t$ to 1, which yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\theta(t) = \omega_R \cos\phi$$

the solution of which is

$$\theta(t) = \omega_R t \cos\phi + \theta_0$$

Let us assume this time that at $t = 0$ $\theta = \pi/2$, i.e., the tip of vector r is on the equator and its longitude is ϕ. Then

$$\theta(t) = \omega_R t \cos\phi + \frac{\pi}{2}$$

First observe that for $\phi = 0$ we get exactly what we had before, θ is going to increase or, in other words, r will continue on its way south at the rate of ω_R per second. On the other hand, if $\phi = \pi$, $\cos\phi = -1$ and then θ is going to *decrease* at the rate of ω_R per second, i.e., r will turn north. So we have now reproduced the basic characteristics of our Rabi spiral. For any other angle ϕ between 0 and π, the progress of r is going to be slowed down. For

$0 < \phi < \pi/2$ $\boldsymbol{r}$ will move south. For $\pi/2 < \phi < \pi$ $\boldsymbol{r}$ will move north instead. But for $\phi = \pi/2$ $\boldsymbol{r}$ will get stuck on the equator.

As was the case with the Larmor precession, wagging $\boldsymbol{r}$ to and fro on the Bloch *Rabi oscillations* sphere with transverse oscillating magnetic fields still does not constitute a mea- *do not* surement. If we start with a fully polarized qubit state, we end with a fully polarized *constitute a* qubit state too. *measurement*

Let us have a look at how probability of finding the qubit in state $|\uparrow\rangle$ varies with time

$$p^0 = \frac{1}{2}\left(1 + r^z\right) = \frac{1}{2}\left(1 + \cos\omega_R t\right)$$

We are again going to make use of the high school trigonometric formula to wrap $\frac{1}{2}(1 + \cos 2\alpha) =$ the above into $\cos^2\alpha$

$$p^0 = \cos^2\frac{\omega_R t}{2} \tag{2.72}$$

At $t = 0$ we have that $p^0 = 1$. Then the probability begins to diminish and at $\omega_R t/2 = \pi/2$ we get $p^0 = 0$, which means that $|\uparrow\rangle$ has flipped completely to $|\downarrow\rangle$. This is the same picture as given by the evolution of $\boldsymbol{r}$, this time expressed in terms of probabilities.

This is also another rare circumstance when an outcome of a quantum mechan- *Exact* ical experiment can be predicted exactly and with certainty for every individual *predictions can* quantum system. If you take a neutron from an $|\uparrow\rangle$ beam and send it into a cham- *be made about* ber with $\boldsymbol{B}_\parallel$ and $\boldsymbol{B}_\perp$ as specified in this section, after $\pi\hbar/(2\mu B_\perp)$ seconds spent in *quantum* the chamber (but not a fraction of a second longer!) the neutron is guaranteed to *systems* be in the $|\downarrow\rangle$ state.[18]

2.10.2   Solution off resonance

Solution (2.65) on page 77 is valid for all values of ω regardless of whether ω is close to ω_L or not. It is valid as long as the approximation we have made, $B_\perp \ll B_\parallel$, is valid. For $\omega = \omega_L$ we found that θ varied linearly with time, eventually swinging onto the other side, i.e., from $\theta = 0$ to $\theta = 180°$.

It is easy to see that for ω very far away from ω_L something quite different is *Solution far* going to happen. Let us recall equation (2.65): *from the* *resonance - the*

$$\theta(t) = \frac{2\mu B_\perp}{\hbar}\frac{\sin\left(\omega_L - \omega\right)t}{\omega_L - \omega} \tag{2.73}$$

qubit does not *absorb energy.*

---

[18]In practice we cannot make this prediction with certainty because the Helmholtz field $\boldsymbol{B}_\perp$ is usually highly non-uniform and because we may not be able to switch it on and off exactly on time. But we want to distinguish here between the fundamental quantum probabilities and probabilities that arise from imperfections of the experimental apparatus.

When $|\omega_L - \omega| >> 0$ the denominator becomes a very large number.[19] Therefore the amplitude of oscillations described by (2.73) is very small and the oscillations themselves are very fast. The result is that vector $\boldsymbol{r}$ keeps pointing up and just vibrates very quickly around $\theta = 0$. The qubit does not absorb energy from the incident radiation.

Solution near the resonance What if $\omega \neq \omega_L$ but they are not so far apart that $\boldsymbol{r}$ gets stuck on $\theta = 0$? Observe that it is only for $\omega = \omega_L$ that θ can wander all over the place. Otherwise $-1 \leq \sin(\omega_L - \omega) t \leq 1$ and θ is restricted to:

$$-\frac{2\mu B_\perp}{\hbar(\omega_L - \omega)} \leq \theta \leq \frac{2\mu B_\perp}{\hbar(\omega_L - \omega)}$$

We can therefore ask the simple question: how far away can we move ω from ω_L so that θ can still reach π – in however circuitous manner? And the answer to this question is

$$\pi = \frac{2\mu B_\perp}{\hbar(\omega_L - \omega)}$$

which yields

$$\omega_L - \omega = \frac{2\mu B_\perp}{\hbar\pi} \tag{2.74}$$

Condition for the effective absorption of energy from the buzzing field $\boldsymbol{B}_\perp$. A somewhat better measure here would be $(\omega_L - \omega)/\omega_L$, because this quantity is non-dimensional and it will let us eliminate $\hbar$ and μ from the equation. Dividing (2.74) by Larmor frequency (given by equation (2.49) on page 73) results in

$$\frac{\omega_L - \omega}{\omega_L} = \frac{2\mu B_\perp}{\hbar\pi}\frac{\hbar}{2\mu B_\parallel} = \frac{B_\perp}{\pi B_\parallel} \tag{2.75}$$

This gives us a measure of how precisely we have to tune the frequency of the buzzing field $\boldsymbol{B}_\perp$ in order to eventuate a qubit flip. Observe that whereas on the one hand the smaller $B_\perp/B_\parallel$ the more accurate is the solution given by equation (2.73), on the other hand, the closer we have to get with ω to ω_L in order to manipulate the qubit according to this equation.

In a typical NMR experiment $B_\parallel$ may be on the order of $12\,\mathrm{T}$. The amplitude of the buzzing field, which is generated by Helmholtz coils, is very tiny, usually on the order of one Gauss, where $1\,\mathrm{Gauss} = 10^{-4}\,\mathrm{T}$.[20] This tells us that we must ensure at least

$$\frac{\omega_L - \omega}{\omega_L} < \frac{10^{-5}}{\pi}$$

---

[19] We should point out that the Planck constant, $\hbar$, in the denominator does not squash $|\omega_L - \omega|$ at all, because there is another $\hbar$ hidden inside μ and they cancel out.

[20] This data is based on the Varian Inova 500 MHz specifications.

2.11   The quantronium

The quantronium [52] is a quantum electronic circuit realization of a qubit. It *Quantronium is* is not a classical electronic device that simulates a quantum system. It is a true *a* quantum system in its own right, even though it does not incorporate obvious *superconducting* microscopic elements such as, e.g., an individual atom of phosphorus embedded *device* in silicon lattice in a precisely defined location and surrounded with controlling electrodes [36] [37]. Instead, the quantronium relies on one of the few macroscopic manifestations of quantum physics, superconductivity. This has the advantage that the circuit can be made using a pretty standard, though not necessarily "industry standard", microelectronic technology.[21]
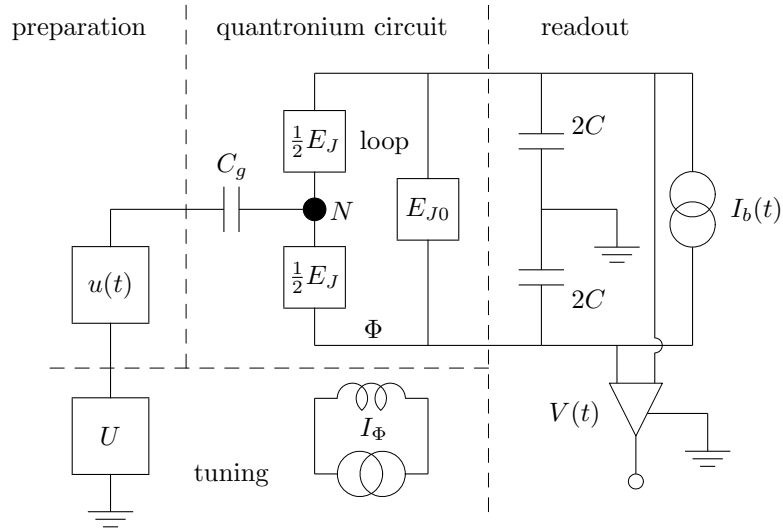
Figure 2.8: The quantronium and its auxiliary electronic circuitry (from [52])

The quantronium circuit together with other auxiliary circuitry is shown in figure 2.8. We cannot analyze the functioning of this circuit in this section in great detail, because this would call for quantum physics background that has not been covered yet. But we will explain enough of it to illustrate how the circuit is used to observe

---

[21]The quantronium was patterned using electron beam lithography and aluminum evaporation. Electron beam lithography is a very precise laboratory technique that allows for nano-level pattern definition. Because of its slowness it cannot be used in mass-produced devices.

the Rabi oscillations and other quantum effects, e.g., the Larmor precession, in the qubit.

Where is the qubit?   The qubit itself is contained in the large black dot to the left of the letter N. The dot symbolizes a low-capacitance superconducting electrode, which in this configuration is called the Cooper pair box. The Cooper pair box is connected to the rest of the circuit through two Josephson tunnel junctions represented in figure 2.8 by the two square boxes with $\frac{1}{2}E_J$ written inside them. The Josephson junctions and the Cooper pair box are biased across the gate capacitance C_g by the voltage source U.

A Cooper pair is a pair of electrons, coupled to each other and to the lattice of the crystal they live in, so that they form a quasiparticle, i.e., a composite and *non-local* quantum object with elementary particle characteristics. The total number of such pairs in the Cooper pair box is N. The energy of the box is quantized, meaning that it can assume several discrete values, which depend on the bias voltage U and the DC current I_Φ that flows in the coil adjacent to the quantronium circuit (shown just below the circuit in figure 2.8) and generates the magnetic field flux Φ in the quantronium circuit loop. The ground energy level and the first excited energy level of the Cooper pair box form a two state quantum system, i.e., a qubit. We associate the ground state with $|\,0\rangle$ and the first excited energy level with $|\,1\rangle$. For the operational parameters of the circuit both states are characterized by the same average number of Cooper pairs in the box in order to make the qubit insensitive to fluctuations of the gate charge.

Tuning the qubit: The bias voltage U and the current I_Φ are used to tune the properties of the qubit with the effect that its Larmor frequency, $\omega_L = 2\mu B_\parallel/\hbar$, is a function of U
$\omega_L = \omega_L(U, I_\Phi)$ and I_Φ. But there is no simple formula that we can use to separate μ from $B_\parallel$. For a given pair of U and I_Φ we get $\mu B_\parallel$ bundled together – although, of course, we can separate them experimentally after we have carried out sufficient measurements on the circuit.

Preparing the qubit with the   The buzzing field $\boldsymbol{B}_\perp$ is represented by the variable voltage $u(t)$ that is super-imposed on U. In order to drive the qubit the frequency of the pulse must match
$u(t)$ pulse   the qubit's Larmor frequency $f_L = \omega_L/(2\pi)$, which, for the circuit drawn in figure 2.8 was 16.4635 ± 0.0008 GHz. By choosing the amplitude and the duration of the pulse, we can swing vector $\boldsymbol{r}$ up and down, thus affecting p^0 and p^1 of the qubit. Having done so we can commence the measurement, which is implemented by the part of the circuit drawn on the right hand side of figure 2.8.

Measuring the qubit   States $|\,0\rangle$ and $|\,1\rangle$ are differentiated by the supercurrent in the loop that develops as the result of the trapezoidal readout pulse $I_b(t)$, which can be sent into the circuit. The current flows through both Josephson junctions labeled by $\frac{1}{2}E_J$ and through

the third Josephson junction on the right hand side of the loop, labeled by E_{J0}. The two capacitors in the readout part of the circuit are meant to reduce phase fluctuations in the loop. The supercurrent in combination with the bias current in the E_{J0} junction can switch the junction to a measurable voltage state $V(t)$. This switching is probabilistic too. There is a high probability that the junction will switch if the qubit is in state $\mid 1\rangle$ and a low probability that it will switch if the qubit is in state $\mid 0\rangle$. So we have to play with two probabilistic processes here: we have quantum probabilities associated with the qubit itself, and then we have another layer of probabilities associated with the readout circuitry. The efficiency of the readout is 60%, i.e., in 60% of cases, the readout circuit will correctly discriminate between $\mid 0\rangle$ and $\mid 1\rangle$. While not perfect this is sufficient to let us observe important quantum phenomena such as Rabi oscillations.

Isolating the qubit

The qubit is isolated from the environment and from the readout circuitry by a variety of means. First, the whole circuit is cooled to $15\,\mathrm{mK}$. Then additional protection is provided by large ratios of E_{J0}/E_J and C/C_J, where C_J is the capacitance of the Josephson junction. Parameters U and I_Φ are chosen so as to eliminate charge fluctuation noise and flux and bias current noise.

Quantronium measurements are an example of sequential exploration of the qubit's statistical ensemble.

The observed Rabi fluctuations of the quantronium are shown in figure 2.9 (A). Every point in this graph is the result of 50,000 identical qubit preparations and measurements, which were carried out in order to collect sufficient statistics for the estimate of the switching probability. The standard deviation on 50,000 measurements of this type is $\sqrt{50,000} \approx 224$, which means that our probability estimates are loaded with no more than 0.5% error.[22]

The buzzing signal $u(t)$ of amplitude $B_\perp = 22\,\mu\mathrm{V}$ and frequency $f_L = 16.4635\,\mathrm{GHz}$ was used for all points. Flux Φ was set to zero.

In order to prepare the qubit for a given point on the graph, the qubit was first allowed to thermalize and align in $U \equiv B_\parallel$, dropping to its ground state $\mid 0\rangle$. Then the qubit was buzzed with $u(t)$ for a specific duration (up to $1\,\mu\mathrm{s}$). As soon as the buzzing had stopped, the trapezoidal pulse $I_b(t)$ was sent into the circuit in order to trigger the switching of the large Josephson junction E_{J0}. If the junction had switched, the resulting pulse $V(t)$ was observed, otherwise there was no pulse. The ratio of "switched" to "not switched" for 50,000 shots/point is what has been plotted in figure 2.9 (A).

Rabi oscillations of quantronium

We can clearly see oscillations in the graph that are the function of the pulse duration. Because the switching probability in the large junction E_{J0} is proportional to p^0, the observed oscillations are indeed Rabi oscillations. For the Larmor fre-

---

[22]See, e.g., chapter 6 in [19].

quency in the tens of GHz range, the Rabi frequency is in the tens of MHz range. The Rabi oscillations for this qubit are therefore about a thousand times slower than the Larmor precession.

Qubit depolarization   The amplitude of the oscillations is clearly damped. This is because the qubit becomes depolarized as it interacts with the environment, however much the circuit's designers had tried to reduce such interaction. This effect is not described by our somewhat simplistic model, nor can it even be described by equations (2.39) and (2.43) on pages 70 and 72 respectively. Nevertheless we can fit the data with an exponentially damped sinusoid and extract the Rabi frequency from it. The fitted curve is drawn in the red.

Dependence of ω_R on $B_\perp$ Having done so, we can repeat the whole experiment for different values of the buzzing signal $u(t)$ amplitude, $B_\perp$, and verify that the Rabi frequency ω_R obtained from the measurements increases linearly with $B_\perp$ as equation (2.68) on page 79 ($\omega_R = 2\mu B_\perp/\hbar$) asserts.

The results of these measurements shown in the right panel on figure 2.9 (A) fully confirm equation (2.68). The Rabi frequency ω_R is indeed directly proportional to $B_\perp$.

The Ramsey experiment   Figure 2.9 (B) shows the result of the Ramsey operation on the quantronium qubit. The operation is carried out as follows. First the qubit is thermalized and brought to state $\mid 0\rangle$. It is then buzzed with $\boldsymbol{B}_\perp$ for the time required to rotate it by $\pi/2$, i.e., by $90°$. The buzzing stops then and the qubit is allowed to precess around the equator for Δt microseconds, whereupon the buzzing is resumed. We buzz the qubit again for the time required to rotate it by the further $90°$ *in normal circumstances*. But recall our discussion of the Ramsey experiment on page 80. Only if the qubit has precessed by a multiple of 2π in time Δt, will it resume it journey south with the same angular polar velocity ω_R. If the qubit has precessed by an odd multiple of π in the time Δt it will instead return back north with the same angular polar velocity ω_R. But for any other angle the qubit has precessed in the time Δt its march south or north will be slowed down by the cosine of the angle. Consequently, when we apply the second buzz to the qubit, it won't be enough to make it go all the way to the top or all the way to the bottom. When the measurement is made the qubit will have some probability of being found in $\mid 1\rangle$ and some probability of being found in $\mid 0\rangle$. This probability will fluctuate with Δt as shown in figure 2.9 (B). The fluctuations observed in this figure, in essence, show us that the qubit indeed precesses with the Larmor frequency around the z direction. The amplitude of the fluctuations diminishes exponentially, as was the case with the Rabi oscillations, even though the qubit is disconnected from the driving force during the time Δt when it is expected to precess freely. The

depolarization observed in this measurement can be used to estimate the so called coherence time of the qubit. Any quantum computations we want to carry out using the qubit, must be completed well before the qubit's quantum state decoheres, which manifests in its depolarization.

We have argued in section 2.10.2 on page 81 that the buzzing signal could force *Resonant* qubit to flip its basis state only if the buzzing frequency was very close to the *absorption by* Larmor frequency of the qubit. Equation (2.75) quantified this by stating that the *the quantronium* qubit could not flip its state at all if $|\omega - \omega_L|/\omega_L > B_\perp/(\pi B_\parallel)$. The insert in figure *qubit* 2.10 (B) shows how the qubit flip probability for the quantronium depends on the frequency of the buzzing signal $u(t)$ for circuit parameters that correspond to the saddle point of the diagram shown in figure 2.10 (A). This saddle point was then used in the Rabi and Ramsey experiments, because of the parametric stability of the circuit in its vicinity.

Every point in the graph in figure 2.10 (B) is the result of 50,000 measurements too. For each measurement the quantronium qubit was first thermalized and allowed to align with $B_\parallel$ in its ground state $|\,0\rangle$. The qubit was then irradiated with microwaves of a given frequency, emitted by a small antenna inserted into the cryostat together with the circuit, for up to 100 ns (T_R in figure 2.9 is about 100 ns) and then the measurement was activated by sending the trapezoidal pulse $I_b(t)$ into the loop of the quantronium circuit. The big Josephson junction E_{J0} would switch sometimes, which would be detected by observing the pulse $V(t)$, and sometimes it wouldn't. Now, whether it switched or not would depend on the state of the qubit, as we have pointed out already, so the switching probability here is related to whether the qubit itself flipped to $|\,1\rangle$ or not. After 50,000 of such trials sufficient statistics were collected to give us the estimate of the switching probability for a given frequency $f = \omega/(2\pi)$. The measurements were repeated while the frequency of the microwave signal was varied so that the whole neighborhood of the resonance point $f = f_L = \omega_L/(2\pi)$ was covered. And this is what is shown in figure 2.10 (B). We can clearly see a well defined peak. The peak can be fitted with the Lorentz absorption curve yielding the exact position of the resonance at $f_L = 16.4635$ GHz and the peak width $\Delta f = 0.8$MHz.

The quantronium is more than just a yet another attempt to fabricate a qubit as a *Quantum* solid state device – and perhaps the first one that gives us more than just a glimmer *systems are* of its quantum nature. It is a very elucidating example of a simple quantum system *probabilistic and* that brings up two of quantum physics' most important features: its probabilistic *extremely fragile* nature and its extreme sensitivity to the environment. Just as importantly, the quantronium illustrates how quantum experiments are carried out in general and how the qubit's statistical ensemble is explored in particular.
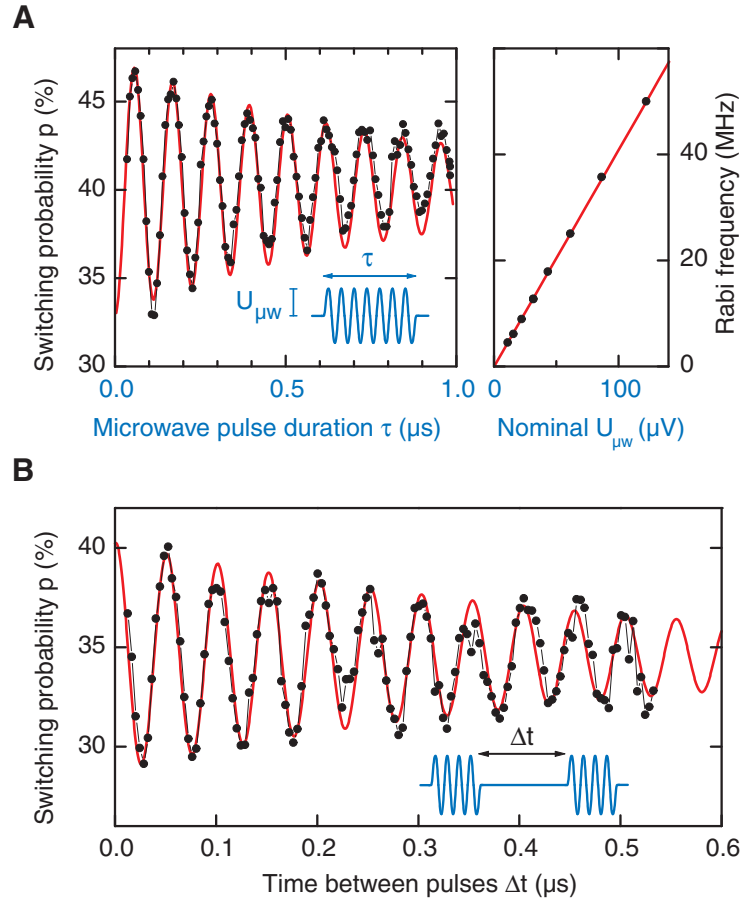
Figure 2.9: Rabi oscillations (A) and Ramsey fringes (B) in quantronium (from [52])
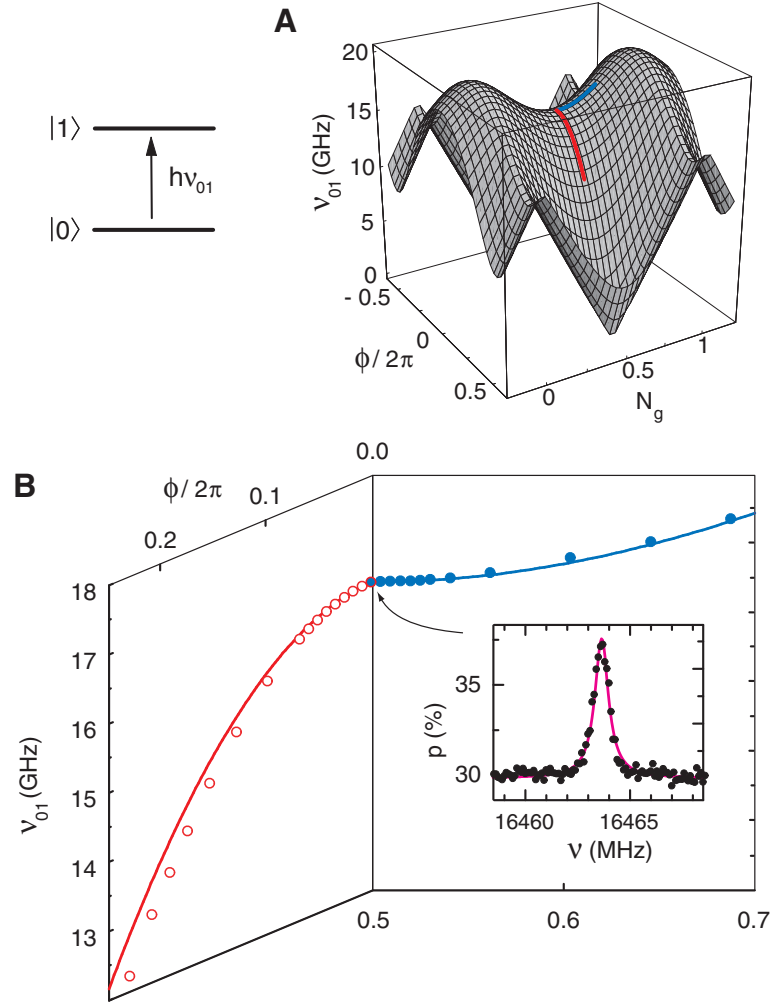
Figure 2.10: Quantronium's transition frequency in function of its various parameters (A). Probability of transition between the $|0\rangle$ and $|1\rangle$ states of the quantronium for circuit parameters that correspond to the saddle point (B) (from [52])

3 Quaternions

The fiducial formalism presented in the previous chapters is inherent to the physics of a qubit: it is complete and entirely physical without any admixture of unnecessary meta-physics. It operates on directly measurable entities such as probabilities and various quantites averaged over a statistical ensemble of a qubit. It accounts for mixtures, fully polarized (pure) states, discrete energy spectrum of a qubit, qubit precession, qubit flipping, and other effects – all within a simple algebra of a conventional probability calculus and without ever having to resort to the use of complex numbers, Hilbert spaces, probability amplitudes and other powerful weapons from the armory of the traditional quantum mechanics. It would make Ernst Mach exuberant.

This may come as a surprise to some physicists, because statements were sometimes made in the past about the impossibility of such a description[1]. Unfortunately quantum mechanics is littered with various statements of this nature, made by some very famous people that – in time – were demonstrated to be blatantly false[2].

Quantum mechanics continues to develop

In spite of quantum mechanics' great maturity, new and surprising results continue to pop up all the time. It was only in 1983 that Berry phase was discovered [6], yet it is such a fundamental quantum effect. It was only in 1996 that a bi-qubit separability criterion was discovered by Peres and Horodeckis [29] [46]. The observation that every quantum system and its evolution can be described entirely in terms of probabilities and various physical parameters averaged over the statistical ensemble of the system (such parameters are called *expectation values*) was published as late as January 2000 [53] by Stefan Weigert of Université de Neuchâtel in Switzerland. The specific observation that such a description can be formulated in terms of a slightly generalized[3] probability calculus is due to Lucien Hardy of The Clarendon Laboratory in Oxford [26] and is even more recent.

In this chapter I am going to demonstrate that the probability calculus for a single qubit maps naturally onto an even simpler type of calculus, for which we will no longer need to bother about vectors and forms, because the state of a qubit as well as its Hamiltonian form will map onto numbers – albeit of a rather special type: quaternions.

---

[1]See, for example, [45] or [20]

[2]See, for example, [4] and [8]

[3]The generalization is confined to the separation between the number of degrees of freedom of the system and the dimension of the system – in classical probability calculus these are the same, in quantum probability calculus the number of degrees of freedom is equal to the square of the dimension of the system.

3.1   Hamilton quaternions

In order to map a vector or a form with four entries onto a number without the loss of information, we must have numbers with four slots. Such numbers were invented by an Irish mathematician, physicist and astronomer William Rowan Hamilton (1805 - 1865) whose accomplishments and insights were so great that the most important mathematical device of quantum physics, the Hamiltonian (we called it the Hamiltonian form), was named after him.

Similarity of quaternions to complex numbers   Quaternions are similar to complex numbers, which have two slots. The slots of a complex number are called *real* and *imaginary*. The imaginary slot is marked by writing the letter i^4 in front of it:

$$z = a + ib$$

Definition of a quaternion   Quaternions have four slots, of which one is real and the remaining three are marked by letters $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$:

$$\boldsymbol{q} = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$$

The coefficients a, b, c, d are normally real, but complex number may be used in their place too. We will see that even though complex numbers are deployed in the mapping between fiducial vectors and forms and quaternions, their use is cosmetic and dictated by tradition rather than necessity. Purely real number based mapping can be used too with almost identical results.

Quaternion commutation relations   The three *imaginary* units of the quaternion world have similar properties to the imaginary unit of complex numbers, i.e.,

$$\boldsymbol{ii} = \boldsymbol{jj} = \boldsymbol{kk} = -1$$

but there is one additional rule:

$$\boldsymbol{ijk} = -1$$

from which the following can be derived:

$$\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}$$
$$\boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}$$
$$\boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}$$

The derivation is quite simple. Consider for example

$$\boldsymbol{ij} = -\boldsymbol{ij}(\boldsymbol{kk}) = -(\boldsymbol{ijk})\boldsymbol{k} = \boldsymbol{k}$$

---

[4]. . . or j if you are an electrical engineer.

Now let us take

$$j(ii)j = -jj = 1 = (ji)(ij) = (ji)k$$

which implies that

$$ji = -k = -ij$$

In other words, quaternion imaginary units do not commute. They anti-commute instead:

$$ij + ji \doteq \{i, j\} = 0$$

and similarly for other pairs. The symbol $\{i, j\}$ is called the anti-commutator.

Quaternion imaginary units anti-commute

3.2   Pauli quaternions

For historic reasons present day physicists do not use symbols i, j and k. Instead they use different symbols[5], namely:

Quaternions and Pauli matrices

$$\sigma_x = ii$$
$$\sigma_y = ij$$
$$\sigma_z = ik$$

where i is the complex numbers i, i.e., $\sqrt{-1}$ and whose resulting properties are:

$$\sigma_x \sigma_x = \sigma_y \sigma_y = \sigma_z \sigma_z = 1$$

and

Pauli matrices commutation and anti-commutation relations

$$\sigma_x \sigma_y = -\sigma_y \sigma_x = (ii)(ij) = -k = iik = i\sigma_z$$
$$\sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x$$
$$\sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y$$

So to avoid a complete disconnect from the way everybody else does things in physics, we're going to use this notation too, even though it is somewhat clumsier than the Hamilton quaternion notation used in mathematics.

Observe that I have used **1** (bold) instead of just 1. This is also due to tradition. People who use σ_x, σ_y and σ_z instead of i, j and k like to think of **1** as an identity matrix rather than a number. But we don't have to do so, neither do we need to look inside σ_x, σ_y and σ_z either, even though these can be represented by matrices as well. To us they will be just 1, i, j and k in disguise.

---

[5] . . . and in effect sometimes do not realize they work with quaternions.

Pauli quaternion commutation rules can be usefully encapsulated into a single expression:

$$\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j = \delta_{ij}\mathbf{1} + i\sum_k \epsilon_{ijk}\boldsymbol{\sigma}_k \qquad (3.1)$$

where i, j and k run through x, y and z, δ_{ij} is the Kronecker delta and ϵ_{ijk} is the Levi Civita tensor, sometimes also called the totally antisymmetric symbol. Indeed, observe that for $i = j$ we get $\delta_{ii} = 1$ but $\epsilon_{iik} = 0$ for every k, and then

$$\boldsymbol{\sigma}_i \boldsymbol{\sigma}_i = \mathbf{1}$$

On the other hand for $i \neq j$ we get $\delta_{ij} = 0$ but $\epsilon_{ijk} = \pm 1$ as long as k is different from both i and j in which case the sign depends on the ordering of whatever i, j and k stand for. In particular if $i = x$ and $j = y$ we get

$$\boldsymbol{\sigma}_x \boldsymbol{\sigma}_y = i\epsilon_{xyz}\boldsymbol{\sigma}_z = i\boldsymbol{\sigma}_z$$

because ϵ_{xyz} is the only non-vanishing ϵ_{xyk} and it is equal to 1.

3.3   From fiducial vectors to quaternions

The mapping between fiducial vectors and quaternions is as simple as it can possibly get. It is going to be linear and we are going to map:

$$
\begin{aligned}
\varsigma_1 &\rightarrow \mathbf{1} \\
\varsigma_x &\rightarrow \boldsymbol{\sigma}_x \\
\varsigma_y &\rightarrow \boldsymbol{\sigma}_y \\
\varsigma_z &\rightarrow \boldsymbol{\sigma}_z
\end{aligned}
$$

and... since there does not have to be a distinction between vectors and forms in the quaternion world, after all they are all just numbers, we're going to map fiducial forms similarly:

$$
\begin{aligned}
\varsigma^1 &\rightarrow \mathbf{1} \\
\varsigma^x &\rightarrow \boldsymbol{\sigma}_x \\
\varsigma^y &\rightarrow \boldsymbol{\sigma}_y \\
\varsigma^z &\rightarrow \boldsymbol{\sigma}_z
\end{aligned}
$$

This way a fiducial vector of a qubit:

$$\boldsymbol{p} = \frac{1}{2}\left(\varsigma_1 + r^x\varsigma_x + r^y\varsigma_y + r^z\varsigma_z\right)$$

becomes

$$\boldsymbol{\rho} = \frac{1}{2}\left(\mathbf{1} + r^x\boldsymbol{\sigma}_x + r^y\boldsymbol{\sigma}_y + r^z\boldsymbol{\sigma}_z\right)$$

and the Hamilton form:

$$\boldsymbol{\eta} = -\mu\left(B_x\boldsymbol{\varsigma}^x + B_y\boldsymbol{\varsigma}^y + B_z\boldsymbol{\varsigma}^z\right)$$

becomes

$$\boldsymbol{H} = -\mu\left(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z\right)$$

We are going to call the two quaternions defined this way, $\boldsymbol{\rho}$ and $\boldsymbol{H}$, a *density quaternion* and a *Hamiltonian quaternion* respectively and ... temporarily. Traditional physics terminology is a little different, because most physicists don't think of these objects as quaternions. They think of them as *operators* instead and call them a *density operator* and a *Hamiltonian operator* or a *Hamiltonian* for short.

Density and Hamiltonian quaternions

Density and Hamiltonian operators

3.4 Expectation values

But how should we map the operation that yields the average energy of the ensemble?

$$\langle\boldsymbol{\eta},\boldsymbol{p}\rangle = -\mu\left(B_x r^x + B_y r^y + B_z r^z\right)$$

The simplest thing to try is to multiply the two quaternions and see what comes out:

$$
\begin{aligned}
\boldsymbol{H\rho} &= -\mu\left(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z\right)\frac{1}{2}\left(\mathbf{1} + r^x\boldsymbol{\sigma}_x + r^y\boldsymbol{\sigma}_y + r^z\boldsymbol{\sigma}_z\right) \\
&= -\frac{\mu}{2}\Big[B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z \\
&\qquad + r^x\left(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z\right)\boldsymbol{\sigma}_x \\
&\qquad + r^y\left(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z\right)\boldsymbol{\sigma}_y \\
&\qquad + r^z\left(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z\right)\boldsymbol{\sigma}_z\Big] \\
&= \ldots
\end{aligned}
$$

Remember that a square of each sigma is $\mathbf{1}$, whereas $\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j$ is some other sigma for $i \neq j$. This makes it easy to collect terms that are proportional to $\mathbf{1}$:

$$
\begin{aligned}
\ldots = -\frac{\mu}{2}\Big[&\left(r^x B_x + r^y B_y + r^z B_z\right)\mathbf{1} \\
&+ \text{various terms multiplied by sigmas}\Big]
\end{aligned}
$$

And so we find our solution $\langle E \rangle = -\mu \left(\boldsymbol{r} \cdot \boldsymbol{B} \right)$ standing right next to $\boldsymbol{1}$. All we need to do is to get rid of the remaining sigma terms. This is easy to do: we

$\Re$ extracts the real part of a quaternion

simply introduce a projection operation that is going to extract a *real* part from the quaternion, similarly to how we extract a real part from a complex number. Define:

$$\Re \left(\boldsymbol{q} \right) = \Re \left(a\boldsymbol{1} + b\boldsymbol{\sigma}_x + c\boldsymbol{\sigma}_y + d\boldsymbol{\sigma}_z \right) \doteq a$$

$\langle E \rangle = 2\Re \left(\boldsymbol{H}\boldsymbol{\rho} \right)$

This way:

$$2\Re \left(\boldsymbol{H}\boldsymbol{\rho} \right) = -\mu \left(\boldsymbol{r} \cdot \boldsymbol{B} \right) = \langle \boldsymbol{\eta}, \boldsymbol{p} \rangle = \langle E \rangle$$

Extracting fiducial vectors from quaternions

In a similar fashion we can obtain probabilities for finding a qubit in a specific state, expressions that are expectation values of a sort too, from the density quaternion $\boldsymbol{\rho}$. Consider the following 4 quaternions:

$$
\begin{aligned}
\boldsymbol{P}^0 &= \frac{1}{2} \left(\boldsymbol{1} + \boldsymbol{\sigma}_z \right) \\
\boldsymbol{P}^1 &= \frac{1}{2} \left(\boldsymbol{1} - \boldsymbol{\sigma}_z \right) \\
\boldsymbol{P}^2 &= \frac{1}{2} \left(\boldsymbol{1} + \boldsymbol{\sigma}_x \right) \\
\boldsymbol{P}^3 &= \frac{1}{2} \left(\boldsymbol{1} + \boldsymbol{\sigma}_y \right)
\end{aligned}
\tag{3.2}
$$

First, observe that they are direct images of the canonical forms in the fiducial space, when expressed in terms of Pauli forms (cf. equations 2.15 on page 64):

$$
\begin{aligned}
\boldsymbol{\omega}^0 &= \frac{1}{2} \left(\varsigma^1 + \varsigma^z \right) \\
\boldsymbol{\omega}^1 &= \frac{1}{2} \left(\varsigma^1 - \varsigma^z \right) \\
\boldsymbol{\omega}^2 &= \frac{1}{2} \left(\varsigma^1 + \varsigma^x \right) \\
\boldsymbol{\omega}^3 &= \frac{1}{2} \left(\varsigma^1 + \varsigma^y \right)
\end{aligned}
$$

Let us apply, say, $\boldsymbol{P}^2$ to $\boldsymbol{\rho}$ and then take $2\Re$ of the result:

$$
\begin{aligned}
\boldsymbol{P}^2 \boldsymbol{\rho} &= \frac{1}{2} \left(\boldsymbol{1} + \boldsymbol{\sigma}_x \right) \frac{1}{2} \left(\boldsymbol{1} + r^x \boldsymbol{\sigma}_x + r^y \boldsymbol{\sigma}_y + r^z \boldsymbol{\sigma}_z \right) \\
&= \frac{1}{4} \left(\boldsymbol{1} + r^x \boldsymbol{\sigma}_x + r^y \boldsymbol{\sigma}_y + r^z \boldsymbol{\sigma}_z + \boldsymbol{\sigma}_x \left(\boldsymbol{1} + r^x \boldsymbol{\sigma}_x + r^y \boldsymbol{\sigma}_y + r^z \boldsymbol{\sigma}_z \right) \right) \\
&= \ldots
\end{aligned}
$$

Now we are again going to collect terms that are proportional to $\mathbf{1}$. There is the single $\mathbf{1}$ in front and then $\boldsymbol{\sigma}_x r^x \boldsymbol{\sigma}_x$ is going to produce another $\mathbf{1}$. All other terms will be multiplied by sigmas. So:

$$\ldots \quad = \quad \frac{1}{4}\left(\mathbf{1} + r^x \mathbf{1}\right)$$
$$+ \text{ various terms multiplied by sigmas}$$

And now taking $2\Re$ of it yields:

$$2\Re\left(\boldsymbol{P}^2 \boldsymbol{\rho}\right) = \frac{1}{2}\left(1 + r^x\right)$$

In the same way one can easily see that:

$$2\Re\left(\boldsymbol{P}^0 \boldsymbol{\rho}\right) = \frac{1}{2}\left(1 + r^z\right)$$
$$2\Re\left(\boldsymbol{P}^1 \boldsymbol{\rho}\right) = \frac{1}{2}\left(1 - r^z\right)$$
$$2\Re\left(\boldsymbol{P}^3 \boldsymbol{\rho}\right) = \frac{1}{2}\left(1 + r^y\right)$$

We have arrived at the following formula that extracts *probabilities* from the density quaternion of a qubit:

$$p^i = 2\Re\left(\boldsymbol{P}^i \boldsymbol{\rho}\right)$$

Probabilities so obtained are consistent with our original mapping

$$\begin{aligned}
\varsigma_1 &\rightarrow \mathbf{1} \\
\varsigma_x &\rightarrow \boldsymbol{\sigma}_x \\
\varsigma_y &\rightarrow \boldsymbol{\sigma}_y \\
\varsigma_z &\rightarrow \boldsymbol{\sigma}_z
\end{aligned}$$

and show that the mapping can be reversed, i.e., in order to extract probabilities from the density quaternion we can either use the $\boldsymbol{P}^i$ quaternions and $2\Re$, or we can simply replace sigmas with varsigmas in the density quaternion and read the probabilities this way.

Generally speaking, as we would represent any measurement on the quantum system by a fiducial form, upon switching to the quaternion representation of the system, we represent any measurement Q on the quantum system by a quaternion $\boldsymbol{Q}$ and the result of this measurement in terms of Q averaged over the ensemble by:

$$\langle Q \rangle = 2\Re\left(\boldsymbol{Q}\boldsymbol{\rho}\right)$$

The formula $p^i = 2\Re\left(\boldsymbol{P}^i \boldsymbol{\rho}\right)$ tells us how to reconstruct the fiducial vector $\boldsymbol{p}$ from the density quaternion $\boldsymbol{\rho}$. Is there an analogous formula that would generate the *form* $\tilde{\boldsymbol{\eta}}$ from the Hamilton quaternion $\boldsymbol{H}$, without resorting to replacing sigmas with the corresponding form-varsigmas?

Extracting fiducial forms from quaternions

Indeed, such a formula can be read out from $\langle \boldsymbol{\eta}, \boldsymbol{p} \rangle$ as follows:

$$\langle E \rangle = \langle \boldsymbol{\eta}, \boldsymbol{p} \rangle = \sum_i \eta_i p^i = \sum_i \eta_i 2\Re\left(\boldsymbol{P}^i \boldsymbol{\rho}\right) = 2\Re\left(\sum_i \eta_i \boldsymbol{P}^i \boldsymbol{\rho}\right) = 2\Re\left(\boldsymbol{H}\boldsymbol{\rho}\right)$$

where we have made use of the fact that $2\Re$ is a linear operation. From this it is now clear that

$$\boldsymbol{H} = \sum_i \eta_i \boldsymbol{P}^i$$

So the way to read coefficients η_i is to express $\boldsymbol{H}$ *not* in terms of sigmas, but in terms of $\boldsymbol{P}^i$ instead!

Let us try it:

$$\begin{aligned}
\boldsymbol{1} &= \boldsymbol{P}^0 + \boldsymbol{P}^1 \\
\boldsymbol{\sigma}_z &= \boldsymbol{P}^0 - \boldsymbol{P}^1 \\
\boldsymbol{\sigma}_x &= 2\boldsymbol{P}^2 - \boldsymbol{1} = 2\boldsymbol{P}^2 - \boldsymbol{P}^0 - \boldsymbol{P}^1 \\
\boldsymbol{\sigma}_y &= 2\boldsymbol{P}^3 - \boldsymbol{1} = 2\boldsymbol{P}^3 - \boldsymbol{P}^0 - \boldsymbol{P}^1
\end{aligned}$$

Substituting this in place of sigmas in

$$\boldsymbol{H} = -\mu(B_x \boldsymbol{\sigma}_x + B_y \boldsymbol{\sigma}_y + B_z \boldsymbol{\sigma}_z)$$

yields

$$\begin{aligned}
\boldsymbol{H} &= -\mu\left(B_x\left(2\boldsymbol{P}^2 - \boldsymbol{P}^0 - \boldsymbol{P}^1\right) + B_y\left(2\boldsymbol{P}^3 - \boldsymbol{P}^0 - \boldsymbol{P}^1\right) + B_z\left(\boldsymbol{P}^0 - \boldsymbol{P}^1\right)\right) \\
&= -\mu\left((B_z - B_x - B_y)\boldsymbol{P}^0 - (B_z + B_x + B_y)\boldsymbol{P}^1 + 2B_x\boldsymbol{P}^2 + 2B_y\boldsymbol{P}^3\right)
\end{aligned}$$

and so we get

$$\begin{aligned}
\eta_0 &= B_z - B_x - B_y \\
\eta_1 &= -B_z - B_x - B_y \\
\eta_2 &= 2B_x \\
\eta_3 &= 2B_y
\end{aligned}$$

which is consistent with $\boldsymbol{\eta} = B_x \boldsymbol{\varsigma}^x + B_y \boldsymbol{\varsigma}^y + B_z \boldsymbol{\varsigma}^z$, where

$$
\begin{aligned}
\boldsymbol{\varsigma}^x &\equiv (-1, -1, 2, 0) \\
\boldsymbol{\varsigma}^y &\equiv (-1, -1, 0, 2) \\
\boldsymbol{\varsigma}^z &\equiv (1, -1, 0, 0)
\end{aligned}
$$

3.5   Mixtures

Now that we know how to switch between quaternions and vectors and forms of the fiducial space of a qubit, let us try to express some ideas we explored in the previous chapter in the pure quaternion language.

The first of these is going to be constructing a mixture of pure states. A pure state is described by a vector $\boldsymbol{r}$ with length 1. Within the fiducial formalism we constructed a mixture by taking a convex linear combination of two or more pure states represented by probability vectors, e.g., $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$:

$$
\boldsymbol{p} = a_1 \boldsymbol{p}_1 + a_2 \boldsymbol{p}_2,
$$
$$
\text{where} \quad a_1 + a_2 = 1
$$

Since the mapping between fiducial vectors and quaternions is linear, the same *Mixing density* should hold for the density quaternion $\boldsymbol{\rho}$: *quaternions*

$$
\begin{aligned}
\boldsymbol{\rho} &= a_1 \boldsymbol{\rho_1} + a_2 \boldsymbol{\rho_2} \\
&= a_1 \frac{1}{2} (\mathbf{1} + r_1^x \boldsymbol{\sigma}_x + r_1^y \boldsymbol{\sigma}_y + r_1^z \boldsymbol{\sigma}_z) + a_2 \frac{1}{2} (\mathbf{1} + r_2^x \boldsymbol{\sigma}_x + r_2^y \boldsymbol{\sigma}_y + r_2^z \boldsymbol{\sigma}_z) \\
&= \frac{1}{2} ((a_1 + a_2)\mathbf{1} + (a_1 r_1^x + a_2 r_2^x)\boldsymbol{\sigma}_x + (a_1 r_1^y + a_2 r_2^y)\boldsymbol{\sigma}_y + (a_1 r_1^z + a_2 r_2^z)\boldsymbol{\sigma}_z)
\end{aligned}
$$

Since $a_1 + a_2 = 1$ the first term in the large brackets, $(a_1 + a_2)\mathbf{1}$ becomes $\mathbf{1}$. The remaining three sigma terms become $r^x \boldsymbol{\sigma}_x + r^y \boldsymbol{\sigma}_y + r^z \boldsymbol{\sigma}_z$, where

$$
\boldsymbol{r} = a_1 \boldsymbol{r}_1 + a_2 \boldsymbol{r}_2
$$

and

$$
\begin{aligned}
\boldsymbol{r} \cdot \boldsymbol{r} &= (a_1 \boldsymbol{r}_1 + a_2 \boldsymbol{r}_2) \cdot (a_1 \boldsymbol{r}_1 + a_2 \boldsymbol{r}_2) \\
&= a_1^2 + a_2^2 + 2 a_1 a_2 \cos(\boldsymbol{r}_1, \boldsymbol{r}_2) \\
&\leq a_1^2 + a_2^2 + 2 a_1 a_2 = (a_1 + a_2)^2 = 1
\end{aligned}
$$

This is the same result we obtained previously for the mixture of two fiducial vectors.

3.6   Qubit evolution

The qubit evolution equation that had the following form in the fiducial formalism

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p} = \frac{1}{2\hbar}\langle \boldsymbol{\eta} \otimes \tilde{\boldsymbol{p}}, \boldsymbol{\epsilon}\rangle$$

and that eventually transformed into an easier readable

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r} = \frac{2\mu}{\hbar}\,\boldsymbol{r} \times \boldsymbol{B}$$

Von Neumann equation   has the following quaternion formulation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\rho} = -\frac{i}{\hbar}\,[\boldsymbol{H}, \boldsymbol{\rho}] \tag{3.3}$$

where

$$[\boldsymbol{H}, \boldsymbol{\rho}] \doteq \boldsymbol{H}\boldsymbol{\rho} - \boldsymbol{\rho}\boldsymbol{H}$$

is called the *commutator* of $\boldsymbol{H}$ and $\boldsymbol{\rho}$. Written in this form the qubit evolution equation is called the Schrödinger-Pauli-von-Neumann equation, or just von Neumann equation for short.

   To transform this equation back into $\mathrm{d}\boldsymbol{p}/\mathrm{d}t = \langle \boldsymbol{\eta} \otimes \tilde{\boldsymbol{p}}, \boldsymbol{\epsilon}\rangle/2\hbar$ calls for a certain amount of tedious algebraic equilibristics, but its transformation into $\mathrm{d}\boldsymbol{r}/\mathrm{d}t = 2\mu\,\boldsymbol{r} \times \boldsymbol{B}/\hbar$ is quite simple.

   First consider the left hand side of the von Neumann equation.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\rho} &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{1} + r^x\boldsymbol{\sigma}_x + r^y\boldsymbol{\sigma}_y + r^z\boldsymbol{\sigma}_z\right) \\
&= \frac{1}{2}\left(\left(\frac{\mathrm{d}}{\mathrm{d}t}r^x\right)\boldsymbol{\sigma}_x + \left(\frac{\mathrm{d}}{\mathrm{d}t}r^y\right)\boldsymbol{\sigma}_y + \left(\frac{\mathrm{d}}{\mathrm{d}t}r^z\right)\boldsymbol{\sigma}_z\right) \\
&= \frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r}\right)\cdot\vec{\boldsymbol{\sigma}}
\end{aligned}$$

because $\mathbf{1}$, $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ are all constants.

The meaning of $\vec{\boldsymbol{\sigma}}$ The symbol $\vec{\boldsymbol{\sigma}}$ is a tri-vector of the three sigmas: $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$. Expressions such as $v^x\boldsymbol{\sigma}_x + v^y\boldsymbol{\sigma}_y + v^z\boldsymbol{\sigma}_z$ occur so often that a typographic shortcut $\boldsymbol{v}\cdot\vec{\boldsymbol{\sigma}}$ was invented. But this should *not* be understood as a real scalar product of two vectors in which infomation is lost. Rather it should be understood as something similar to $v^x\boldsymbol{e}_x + v^y\boldsymbol{e}_y + v^z\boldsymbol{e}_z \doteq \boldsymbol{v}\cdot\vec{\boldsymbol{e}}$. Here no information is lost as the operation is performed. There are three fully extractable components of $\boldsymbol{v}$ on both sides of the equation.

Now let us turn to the commutator:

$$\begin{aligned}
\boldsymbol{H}\boldsymbol{\rho} &- \boldsymbol{\rho}\boldsymbol{H} \\
&= -\frac{\mu}{2}\Big[(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z)(\mathbf{1} + r^x\boldsymbol{\sigma}_x + r^y\boldsymbol{\sigma}_y + r^z\boldsymbol{\sigma}_z) \\
&\quad - (\mathbf{1} + r^x\boldsymbol{\sigma}_x + r^y\boldsymbol{\sigma}_y + r^z\boldsymbol{\sigma}_z)(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z)\Big]
\end{aligned}$$

$$\ldots$$

Before we plunge into the fury of computation, let us have a sanguine look at the equation first. To begin with we are going to have terms resulting from multiplication of the Hamilton quaternion by the $\mathbf{1}$ of the density quaternion. These will produce $\boldsymbol{B} \cdot \vec{\boldsymbol{\sigma}}$ from $\boldsymbol{H}\boldsymbol{\rho}$ and $-\boldsymbol{B} \cdot \vec{\boldsymbol{\sigma}}$ from $-\boldsymbol{\rho}\boldsymbol{H}$, so they'll cancel in effect. Next we're going to have terms resulting from $\boldsymbol{\sigma}_i$, $i = x, y, z$ multiplying itself, i.e., $\boldsymbol{\sigma}_x\boldsymbol{\sigma}_x$, etc. For each sigma $\boldsymbol{\sigma}_i\boldsymbol{\sigma}_i = \mathbf{1}$, so these terms are going to produce $\boldsymbol{B} \cdot \boldsymbol{r}\,\mathbf{1}$ from $\boldsymbol{H}\boldsymbol{\rho}$, which is exactly what we had in $2\Re(\boldsymbol{H}\boldsymbol{\rho})$, and $-\boldsymbol{B} \cdot \boldsymbol{r}\,\mathbf{1}$ from $\boldsymbol{\rho}\boldsymbol{H}$, so they'll cancel too.

The only terms that are going to survive then will be the asymmetric terms:

$$\begin{aligned}
\ldots \;=\; -\frac{\mu}{2}\Big[&\big(B_x r^y \boldsymbol{\sigma}_x\boldsymbol{\sigma}_y + B_x r^z \boldsymbol{\sigma}_x\boldsymbol{\sigma}_z \\
&+ B_y r^x \boldsymbol{\sigma}_y\boldsymbol{\sigma}_x + B_y r^z \boldsymbol{\sigma}_y\boldsymbol{\sigma}_z \\
&+ B_z r^x \boldsymbol{\sigma}_z\boldsymbol{\sigma}_x + B_z r^y \boldsymbol{\sigma}_z\boldsymbol{\sigma}_y\big) \\
&- \big(B_x r^y \boldsymbol{\sigma}_y\boldsymbol{\sigma}_x + B_x r^z \boldsymbol{\sigma}_z\boldsymbol{\sigma}_x \\
&+ B_y r^x \boldsymbol{\sigma}_x\boldsymbol{\sigma}_y + B_y r^z \boldsymbol{\sigma}_z\boldsymbol{\sigma}_y \\
&+ B_z r^x \boldsymbol{\sigma}_x\boldsymbol{\sigma}_z + B_z r^y \boldsymbol{\sigma}_y\boldsymbol{\sigma}_z\big)\Big] \\
=\; \ldots
\end{aligned}$$

Observe that the expression in the second round bracket is the same as the expression in the first round bracket but with sigmas ordered the other way. Switching them around produces minus, which cancels with the minus of the commutator, so that we end up with:

$$\begin{aligned}
\ldots \;=\; -\frac{\mu}{2}2\big(&B_x r^y \boldsymbol{\sigma}_x\boldsymbol{\sigma}_y + B_x r^z \boldsymbol{\sigma}_x\boldsymbol{\sigma}_z \\
&+ B_y r^x \boldsymbol{\sigma}_y\boldsymbol{\sigma}_x + B_y r^z \boldsymbol{\sigma}_y\boldsymbol{\sigma}_z \\
&+ B_z r^x \boldsymbol{\sigma}_z\boldsymbol{\sigma}_x + B_z r^y \boldsymbol{\sigma}_z\boldsymbol{\sigma}_y\big) \\
=\; \ldots
\end{aligned}$$

But here in this expression we also have sigma pairs that are just switched around. Reordering them "the right way" will produce minuses, so that we'll end up with

$$\ldots \quad = \quad -\mu\Big[\left(B_x r^y - B_y r^x\right)\boldsymbol{\sigma}_x\boldsymbol{\sigma}_y$$
$$+\left(B_y r^z - B_z r^y\right)\boldsymbol{\sigma}_y\boldsymbol{\sigma}_z$$
$$+\left(B_z r^x - B_x r^z\right)\boldsymbol{\sigma}_z\boldsymbol{\sigma}_x\Big]$$
$$= \quad \ldots$$

Finally, let us replace $\boldsymbol{\sigma}_x\boldsymbol{\sigma}_y$ with $i\boldsymbol{\sigma}_z$, and similarly for the other two pairs to get:

$$\ldots \quad = \quad -i\mu\Big[\left(B_x r^y - B_y r^x\right)\boldsymbol{\sigma}_z + \left(B_y r^z - B_z r^y\right)\boldsymbol{\sigma}_x + \left(B_z r^x - B_x r^z\right)\boldsymbol{\sigma}_y\Big]$$
$$= \quad \ldots$$

Observe that the expressions in the round brackets are components of a vector product $\boldsymbol{B} \times \boldsymbol{r}$. We can therefore rewrite our result in a more compact form as:

$$[\boldsymbol{H},\boldsymbol{\rho}] = -i\mu\left(\boldsymbol{B} \times \boldsymbol{r}\right)\cdot\vec{\boldsymbol{\sigma}}$$

And so the von Neumann equation for the quaternion-described qubit turns into:

$$\frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r}\right)\cdot\vec{\boldsymbol{\sigma}} = -\frac{i}{\hbar}\left(-i\mu\right)\left(\boldsymbol{B} \times \boldsymbol{r}\right)\cdot\vec{\boldsymbol{\sigma}}$$

which then yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{r} = \frac{2\mu}{\hbar}\,\boldsymbol{r} \times \boldsymbol{B}$$

3.7   Why does it work?

Before I attempt to answer this question, let me first explain that all that we've done with the *sigma*-quaternions used by the physicists we could have done just as easily with the original Hamilton $(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$-quaternions. The imaginary unit, i, that pops up in the von-Neumann equation is there only because of the way we defined the sigmas as i times Hamilton quaternions $\boldsymbol{i}$, $\boldsymbol{j}$ or $\boldsymbol{k}$ and does not really represent anything deep or fundamental. It is merely *ornamental* and rather misleading.

There is no need for $i = \sqrt{-1}$ when using Hamilton quaternions

If we were to use the original Hamilton quaternions, our equations would look very similar with only a sign different here or there and without any imaginary units. For example, if we were to map:

$$\varsigma_1 \quad \leftrightarrow \quad 1$$

$$\begin{aligned}
\varsigma_x &\leftrightarrow \boldsymbol{i} \\
\varsigma_y &\leftrightarrow \boldsymbol{j} \\
\varsigma_z &\leftrightarrow \boldsymbol{k}
\end{aligned}$$

we would get:

$$\langle \boldsymbol{\eta}, \boldsymbol{p} \rangle = -2\Re\left(\boldsymbol{H}\boldsymbol{\rho}\right)$$

But we could be a little fancier here and we could, for example, map forms onto *conjugate* quaternions, e.g.,

$$\boldsymbol{\eta} = B_x\varsigma^x + B_y\varsigma^y + B_z\varsigma^z \quad \rightarrow \quad -B_x\boldsymbol{i} - B_y\boldsymbol{j} - B_z\boldsymbol{k}$$

Forms could be mapped on conjugate quaternions

and this would yield

$$\langle E \rangle = 2\Re\left(\boldsymbol{H}\boldsymbol{\rho}\right)$$

The von-Neumann equation in the Hamilton quaternion formalism and with forms mapped onto normal quaternions, not the conjugate ones, is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\rho} = \frac{1}{\hbar}\left[\boldsymbol{H}, \boldsymbol{\rho}\right]$$

there is no i here and no minus either. The minus could be restored by mapping fiducial forms onto conjugate quaternions instead.

So, whether we use sigmas or $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ we get much the same picture - and we still don't have to do our quantum mechanics with complex numbers.

Now, why does the quaternion trick work? The reason for this is that quaternions are general and powerful and their peculiar commutation properties encode both a scalar product (or a *dot* product as some call it) and a vector product. Consider this:

$$\begin{aligned}
\boldsymbol{ab} &= (a^x\boldsymbol{i} + a^y\boldsymbol{j} + a^z\boldsymbol{k})(b^x\boldsymbol{i} + b^y\boldsymbol{j} + b^z\boldsymbol{k}) \\
&= (a^xb^x + a^yb^y + a^zb^z)(-1) \\
&\quad + [(a^yb^z - a^zb^y)\boldsymbol{jk} + (a^zb^x - a^xb^z)\boldsymbol{ki} + (a^xb^y - a^yb^x)\boldsymbol{ij}] \\
&= -\vec{a}\cdot\vec{b} + \left(\vec{a}\times\vec{b}\right)^x\boldsymbol{i} + \left(\vec{a}\times\vec{b}\right)^y\boldsymbol{j} + \left(\vec{a}\times\vec{b}\right)^z\boldsymbol{k}
\end{aligned}$$

So this is where all the magic comes from. To extract the scalar product from the quaternion we can either just take the real part of it, the $\Re$, and multiply by whatever coefficient is needed to get the right answer, or we could take $\boldsymbol{ab} + \boldsymbol{ba}$ and then the $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$-part of the quaternion would cancel out. To extract the vector product from the quaternion we could simply extract the coefficients that multiply $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ or we could take $\boldsymbol{ab} - \boldsymbol{ba}$ and then the real part of the quaternion, the scalar product part, would cancel out.

Quaternions encode a dot and a cross product of two vectors in a single expression

To extract the coefficients that multiply i, j and k we can use operators similar to $\Re$. A complex number operator that extracts the imaginary component of a complex number is called $\Im$. But in case of the quaternions we need three such operators, and so we're going to call them $\Im_i$, $\Im_j$, and $\Im_k$.

Any three-dimensional theory can be mapped onto quaternions

Consequently any theory that contains a combination of scalar and vector products can be mapped onto quaternions and special quaternion rules devised to extract whatever equations of the original theory are wanted.

Quaternions simply encode elementary 3-dimensional vector algebra in the form of "numbers" with 4 slots each.

Quaternions were invented to describe rotations

For this reason Hamilton was able to encode rotations and their combinations by using quaternions – this is, in fact, what he invented them for. Various equations of special and even general relativity can be mapped onto quaternion algebra too.

Our theory of qubits that describes their statistical ensembles in terms of probability vectors eventually resolves to scalar and vector products, such as $\langle E \rangle = -\mu \left(\boldsymbol{r} \cdot \boldsymbol{B} \right)$ and $\mathrm{d}\boldsymbol{r}/\mathrm{d}t = 2\mu \left(\boldsymbol{r} \times \boldsymbol{B} \right)/\hbar$ – and so, it is only natural that a mapping should exist that lets us express it in terms of quaternions.

That a mapping comes out to be so simple, with Pauli vectors and forms mapping directly onto 1 (or $\boldsymbol{1}$) and $(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$ (or $(\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z)$) derives from the fact that I *had* defined ς_1, ς_x, ς_y and ς_z so that this would be the case – knowing in advance the result I wanted to achieve.

But I didn't cheat. $\boldsymbol{p} = \frac{1}{2} \left(\varsigma_1 + r^x \varsigma_x + r^y \varsigma_y + r^z \varsigma_z \right)$ indeed fully describes a qubit state, meaning a state of a statistical ensemble that represents it.

A more profound question should be, why

$$\boldsymbol{p} = \frac{1}{2} \begin{pmatrix} 1 + r^z \\ 1 - r^z \\ 1 + r^x \\ 1 + r^y \end{pmatrix}$$

works.

The reason why this works is because it is so general.

A fiducial vector of a qubit is the most general probability vector with 4 slots and with the first two slots normalized.

The first two entries in $\boldsymbol{p}$ represent the fact that a beam of qubits splits in two in the presence of a magnetic field that has a non-vanishing gradient in the z direction. Individual qubits must go either *up* or *down*. *Tertium non datur.* This exhausts all possibilities. Consequently $p^0 + p^1 = 1$. To express p^0 and p^1 as $(1 + r^z)/2$ and $(1 - r^z)/2$ with $r^z \in [0, 1]$ merely parametrizes this observation. But the experiment also tells us that the description of a qubit in terms of r^z alone is incomplete. Knowing p^0 and p^1 is *not* enough to fully describe the state of a

qubit. In order to fully describe the qubit we must know how it is going to behave in magnetic fields whose gradients are in the x and y directions too. Hence the additional two terms, which are similarly parametrized by r^x and r^y.

Contrary to what some physicists think and say, a qubit can be polarized in *any* direction in the 3-dimensional space. This can be always *confirmed* by moving a magnet in all possible directions until the beam of qubits no longer splits and all qubits in it are deflected in the same direction – that corresponds to $\boldsymbol{r}$. Such a direction can be always found, if a beam is fully polarized. If such a direction does not exist, the beam is not fully polarized. It is a mixture.

So we come to a surprising conclusion: the reason why $\boldsymbol{p}$ describes a qubit is simply because it has enough slots to do so. The reason why quaternions describe a qubit is because they are general and flexible enough to fit any theory that has scalar and vector products in it.

An astute reader may stop me here and ask: "OK about $\boldsymbol{p}$ and quaternions, but what about $\mathrm{d}\boldsymbol{r}/\mathrm{d}t = 2\mu\left(\boldsymbol{r}\times\boldsymbol{B}\right)/\hbar$ and $\langle E\rangle = -\mu\,\boldsymbol{r}\cdot\boldsymbol{B}$? Why do these work?"

This is indeed where the real physics is. What these equations say is that a qubit behaves *on average* like a classical magnetic dipole. This *on average* behavior represents the *classical* thermodynamic limit of the theory. If you immerse a macroscopic sample full of qubits in a magnetic field $\boldsymbol{B}$, you will be able to observe a precession of the *average* magnetic field of the sample exactly as described by $\mathrm{d}\boldsymbol{r}/\mathrm{d}t = 2\mu\left(\boldsymbol{r}\times\boldsymbol{B}\right)/\hbar$. You will be able to observe other effects we discussed in the previous sections too.

A qubit behaves like a classical magnetic dipole on average.

This, in quick summary, is *all* that quantum mechanics can say about a qubit: that a beam of qubits splits in two in the presence of varying magnetic fields, and that an ensemble of identically prepared fully polarized qubits behaves like a classical magnetic dipole *on average*.

Quantum mechanics is a probability theory. All its pronouncements refer only to statistical ensembles of quantum objects. Quantum mechanics has nothing whatsoever to say about a particular individual quantum object – unless in very special and rare circumstances when a behavior of the whole ensemble can be predicted with 100% certainty. In this case only we can be certain about what an individual quantum object is going to do.

4 The Unitary Formalism

The so called *unitary* description of qubits is what arises when the $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ quaternions are "unpacked" into matrices.

> Why should we bother about unpacking them in the first place, if we can obtain all the information we need directly from the quaternion picture, or by using fiducial vectors and forms? This is indeed a very good question. It translates into another even more profound question. Since the fiducial formalism and its mapping onto quaternions – otherwise known as the density operator formalism – already cover all qubit physics, what new physics can we possibly arrive at by unpacking the sigmas? Are we merely going to delude ourselves with unnecessary metaphysics? This question latches directly onto the business of quantum computing, because the whole idea indeed derives from the *notation* of the unitary calculus.

What do we mean by "unpacking the sigmas"? We are going to look for matrix representations of sigmas that have the same commutation and anti-commutation properties, namely:

Matrix representation of quaternions

$$\boldsymbol{\sigma}_x^2 = \boldsymbol{\sigma}_y^2 = \boldsymbol{\sigma}_z^2 = \mathbf{1}$$
$$\boldsymbol{\sigma}_x \boldsymbol{\sigma}_y \boldsymbol{\sigma}_z = i\mathbf{1}$$

The other sigma properties, e.g., $\boldsymbol{\sigma}_x \boldsymbol{\sigma}_y = -\boldsymbol{\sigma}_y \boldsymbol{\sigma}_x = i\boldsymbol{\sigma}_z$, can be derived from the two, as we have seen done with the Hamilton quaternions $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ – and we are going to help ourselves with these too.

4.1 Pauli matrices

It is easy to see that we cannot represent the sigmas by numbers alone, real or complex, because the sigmas anti-commute and neither real nor complex numbers do. But matrices do not commute in general either and so the simplest representation of sigmas can be sought in the form of 2×2 matrices. Because we have the i, the imaginary unit, in our commutation relations for sigmas, we will have to consider 2×2 complex matrices. Pure real matrices will no longer do, because how would we generate the i. Also it will become clear, after we will have completed the exercise, that if we were to use the Hamilton's original $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$ symbols instead, we would still end up with complex matrices.

Let us begin with the following general parametrization of 2×2 matrices

$$\boldsymbol{\sigma}_x = \left(\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right)$$

$$\boldsymbol{\sigma}_y = \left(\begin{array}{cc} b_{11} & b_{12} \\ b_{21} & b_{22} \end{array} \right)$$

$$\boldsymbol{\sigma}_z = \left(\begin{array}{cc} c_{11} & c_{12} \\ c_{21} & c_{22} \end{array} \right)$$

where a_{ij}, b_{ij} and c_{ij} are some complex numbers.

The first rule that applies equally to all sigmas is that they square to $\mathbf{1}$. Let us perform the corresponding computation on $\boldsymbol{\sigma}_z$:

$$
\begin{aligned}
\boldsymbol{\sigma}_z \boldsymbol{\sigma}_z &= \left(\begin{array}{cc} c_{11} & c_{12} \\ c_{21} & c_{22} \end{array} \right) \left(\begin{array}{cc} c_{11} & c_{12} \\ c_{21} & c_{22} \end{array} \right) \\
&= \left(\begin{array}{cc} c_{11}^2 + c_{12}c_{21} & c_{12}(c_{11} + c_{22}) \\ c_{21}(c_{11} + c_{22}) & c_{22}^2 + c_{12}c_{21} \end{array} \right) \\
&= \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right)
\end{aligned}
$$

This yields the following two groups of equations:

$$
\begin{aligned}
c_{12}(c_{11} + c_{22}) &= 0 \\
c_{21}(c_{11} + c_{22}) &= 0
\end{aligned}
$$

and

$$
\begin{aligned}
c_{11}^2 + c_{12}c_{21} &= 1 \\
c_{22}^2 + c_{12}c_{21} &= 1
\end{aligned}
$$

The first two equations can be satisfied by *either*

$$c_{11} = -c_{22}$$

or by

$$c_{12} = c_{21} = 0$$

Suppose the latter holds. In this case we get that

$$c_{11}^2 = c_{22}^2 = 1$$

which implies that
$$c_{11} = \pm 1 = c_{22}$$

We therefore end up with two possibilities:

$$\boldsymbol{\sigma}_z = \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{or}$$

$$\boldsymbol{\sigma}_z = \pm \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The first solution is just $\mathbf{1}$. This is not a good solution here, because we want $\boldsymbol{\sigma}_z$ to be something other than $\mathbf{1}$, so we're going to take the second one, and we're going to choose the plus sign for it:

$$\boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Observe that in this case we ended up with *both*

$$c_{12} = c_{21} = 0$$

and

$$c_{11} = -c_{22}$$

Even though we were prepared to be more general. The other solution in this group, resulting from the generality, was $\mathbf{1}$.

For the remaining two matrices, $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_y$ we have to choose the other option, i.e.,

$$a_{12} \neq 0 \quad \text{and}$$
$$a_{21} \neq 0$$

and similarly for $\boldsymbol{\sigma}_y$ and the bs, because otherwise we'd end up with either $\mathbf{1}$ or $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ again. But if $a_{12} \neq 0$ and $a_{21} \neq 0$ then we must have

$$a_{11} = -a_{22} \doteq a \quad \text{and}$$
$$b_{11} = -b_{22} \doteq b$$

where a and b are such that

$$a^2 + a_{12}a_{21} = 1 \quad \text{and}$$
$$b^2 + b_{12}b_{21} = 1$$

Now let us make use of the anti-commutation rule $\boldsymbol{\sigma}_x\boldsymbol{\sigma}_z + \boldsymbol{\sigma}_z\boldsymbol{\sigma}_x = \mathbf{0}$. First we have:

$$\boldsymbol{\sigma}_x\boldsymbol{\sigma}_z = \begin{pmatrix} a & a_{12} \\ a_{21} & -a \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} a & -a_{12} \\ a_{21} & a \end{pmatrix}$$

but

$$\boldsymbol{\sigma}_z\boldsymbol{\sigma}_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} a & a_{12} \\ a_{21} & -a \end{pmatrix}$$

$$= \begin{pmatrix} a & a_{12} \\ -a_{21} & a \end{pmatrix}$$

Adding these two yields

$$\boldsymbol{\sigma}_x\boldsymbol{\sigma}_z + \boldsymbol{\sigma}_z\boldsymbol{\sigma}_x = 2\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$$

For the anti-commutation rule to hold a must be zero. Since the same argument is going to work for $\boldsymbol{\sigma}_y$ as well we get

$$a = 0 \quad \text{and}$$
$$b = 0$$

In summary:

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & a_{12} \\ a_{21} & 0 \end{pmatrix}$$

$$\boldsymbol{\sigma}_y = \begin{pmatrix} 0 & b_{12} \\ b_{21} & 0 \end{pmatrix}$$

But recall that as a and b vanish we are left with

$$a_{12}a_{21} = 1 \quad \text{and}$$
$$b_{12}b_{21} = 1$$

We can easily satisfy these equations by setting $a_{12} = x$ and $a_{21} = 1/x$ and similarly $b_{12} = y$ and $b_{21} = 1/y$, so that

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & x \\ x^{-1} & 0 \end{pmatrix}$$

$$\boldsymbol{\sigma}_y = \begin{pmatrix} 0 & y \\ y^{-1} & 0 \end{pmatrix}$$

Now we are ready to make use of the rule stating that

$$\boldsymbol{\sigma}_x \boldsymbol{\sigma}_y \boldsymbol{\sigma}_z = i\mathbf{1}$$

Substituting our matrix expressions for $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ yields:

$$\begin{pmatrix} 0 & x \\ x^{-1} & 0 \end{pmatrix} \begin{pmatrix} 0 & y \\ y^{-1} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$= \begin{pmatrix} x/y & 0 \\ 0 & -y/x \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$$

This yields the following 2 equations in combination with what we have arrived at already:

$$\begin{aligned} x/y &= i \\ -y/x &= i \end{aligned}$$

both of which solve to

$$y = -ix$$

Since $1/y = 1/(-ix) = ix^{-1}$ we end up with

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & x \\ x^{-1} & 0 \end{pmatrix} \tag{4.1}$$

$$\boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -ix \\ ix^{-1} & 0 \end{pmatrix} \tag{4.2}$$

$$\boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

However surprising this may be to some physicists on account of the x factor it is very easy to check, e.g., with Maple or Mathematica or manually even, that these matrices indeed satisfy all quaternion relations expected of $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$, which is sufficient to get the right expressions for $\langle E \rangle = -\mu \left(\boldsymbol{r} \cdot \boldsymbol{B} \right)$ via the $\langle E \rangle = 2\Re \left(\boldsymbol{H} \boldsymbol{\rho} \right)$ and for $\mathrm{d}\boldsymbol{r}/\mathrm{d}t = 2\mu \left(\mathrm{r} \times \mathrm{B} \right)/\hbar$ via the $\mathrm{d}\boldsymbol{\rho}/\mathrm{d}t = -i \left[\boldsymbol{H}, \boldsymbol{\rho} \right]/\hbar$.

The choice of $x = 1$ is a natural one in this context – though not strictly necessary at this stage. But it makes $\boldsymbol{\sigma}_x$ pleasingly symmetric and $\boldsymbol{\sigma}_y$ pleasingly Hermitian, i.e., such that $A_{ij} = A_{ji}^*$ and both properties will prove useful as we go along.

Hermitian matrices are of special importance in quantum physics. We will see on page 125 that x will be further restricted to $e^{i\phi_x}$, so that $\boldsymbol{\sigma}_y$, even with this remaining degree of freedom left, will end up being Hermitian anyway –

Hermitian matrices

as will $\boldsymbol{\sigma}_x$. The condition $A_{ij} = A_{ji}^*$ says that if you transpose a Hermitian matrix $\boldsymbol{A}$ and then complex conjugate it, the resulting matrix is the same as the original one:

$$\left(\boldsymbol{A}^T \right)^* = \boldsymbol{A}$$

Hermitian conjugation

There is a special symbol, † that represents this, so called, *Hermitian conjugation*:

$$\boldsymbol{A}^\dagger \doteq \left(\boldsymbol{A}^T \right)^* \tag{4.3}$$

Using this symbol we can say that operator $\boldsymbol{A}$ is Hermitian when

$$\boldsymbol{A}^\dagger = \boldsymbol{A}$$

Pauli matrices

Having made this choice, and we are *not* going to forget about it, instead working on its further justification as we develop the unitary formalism, we arrive at the quaternion representation in terms of Pauli matrices:

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.4}$$

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{4.5}$$

$$\boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{4.6}$$

$$\boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.7}$$

It is useful to express the canonical basis in the space of 2×2 matrices in terms of Pauli matrices – a procedure somewhat similar to what we did earlier with canonical forms in the fiducial space and Pauli forms (cf. equations 2.15 on page 64). And so:

$$\boldsymbol{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2} \left(\mathbf{1} + \boldsymbol{\sigma}_z \right)$$

$$\boldsymbol{M}_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2} \left(\boldsymbol{\sigma}_x + i\boldsymbol{\sigma}_y \right)$$

$$\boldsymbol{M}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \frac{1}{2} \left(\boldsymbol{\sigma}_x - i\boldsymbol{\sigma}_y \right)$$

$$\boldsymbol{M}_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} \left(\mathbf{1} - \boldsymbol{\sigma}_z \right) \tag{4.8}$$

Note that only M_0 and M_3 bear similarity to P^0 and P^1 of equations 3.2 on page 96. Matrix representations of P^2 and P^3 are

$$P^2 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$P^3 = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$$

Observe that all three sigmas are *traceless*, i.e., such that the sum of their diag- *Pauli matrices* onal elements is zero. The only matrix here that has a non-vanishing trace of 2 is *are traceless* **1**. We can therefore use the matrix operation of taking trace, Tr, in place of the quaternion operation $\Re$. Also note that since $\text{Tr}\,(\mathbf{1}) = 2$ we can drop the 2 factor that appeared in front of $\Re$, i.e.,

$$2\Re\,(\boldsymbol{H}\boldsymbol{\rho}) = \text{Tr}\,(\boldsymbol{H}\boldsymbol{\rho})$$

It is this $\text{Tr}\,(\mathbf{1}) = 2$ in fact that is responsible for some of the 1/2s and 2s that appear in definitions of our probability vector and its close cousin, the density quaternion, as well as in the definitions of Pauli forms and of the metric tensor of the fiducial space.

Of the three Pauli matrices one is imaginary, the $\boldsymbol{\sigma}_y$. Could we do away with imaginary numbers if we were to carry out the procedure for Hamilton quaternions $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$? The answer is "no", we wouldn't do better. Recall that

$$\boldsymbol{i} = -i\boldsymbol{\sigma}_x$$
$$\boldsymbol{j} = -i\boldsymbol{\sigma}_y$$
$$\boldsymbol{k} = -i\boldsymbol{\sigma}_z$$

This transformation would make $\boldsymbol{j}$ real, but then we'd end up with imaginary $\boldsymbol{i}$ and $\boldsymbol{k}$.

If we were to substitute a_{ij}, b_{ij} and c_{ij}, as we did previously, for $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$, equation $\boldsymbol{k}\boldsymbol{k} = -\mathbf{1}$ would yield:

$$c_{12}\,(c_{11} + c_{22}) = 0$$
$$c_{21}\,(c_{11} + c_{22}) = 0$$

as before, but this time

$$c_{11}^2 + c_{12}c_{21} = -1$$
$$c_{22}^2 + c_{12}c_{21} = -1$$

In case $c_{12} = c_{21} = 0$ we get that $c_{11}^2 = -1$ and $c_{22}^2 = -1$, which implies $c_{11} = \pm i = c_{22}$ right away.

There is no real
2×2
representation
of quaternions

In summary, whereas we can describe qubit dynamics and kinematics *completely* in terms of real numbers and measurable probabilities alone, as long as we work either within the fiducial formalism or within the density quaternion framework, the moment we "unpack" the quaternions we have to let imaginary numbers through this back-door.

Like Alice falling into the rabbit hole, we are going to encounter some very strange creatures and notions, some of which may well belong in the fantasy world.

4.2 The basis vectors and the Hilbert space

Density matrix

A fully unpacked density quaternion of a qubit becomes a 2×2 complex matrix that looks as follows

$$\boldsymbol{\rho} = \frac{1}{2}\left(\mathbf{1} + r^x\boldsymbol{\sigma}_x + r^y\boldsymbol{\sigma}_y + r^z\boldsymbol{\sigma}_z\right) = \frac{1}{2}\begin{pmatrix} 1 + r^z & r^x - ir^y \\ r^x + ir^y & 1 - r^z \end{pmatrix} \qquad (4.9)$$

and in this form it is called a *density operator* or a *density matrix* of a qubit.

We had seen an operator $\boldsymbol{P}$ in section 1.8 (page 32) that talked about "Transformations of mixtures" represented as a sum of tensor products of basis vectors and forms, where the vectors and forms were ordered "the other way round" so that they wouldn't eat each other:

$$\boldsymbol{P} \doteq \sum_j \sum_i p^j{}_i \boldsymbol{e}_j \otimes \boldsymbol{\omega}^i$$

This trick can be applied to every linear operator that is represented by a matrix, including $\boldsymbol{\rho}$, and in this case we can write:

$$\begin{aligned} \boldsymbol{\rho} &= \frac{1}{2}\Big((1 + r^z)\,\boldsymbol{e}_0 \otimes \boldsymbol{\omega}^0 + (r^x - ir^y)\,\boldsymbol{e}_0 \otimes \boldsymbol{\omega}^1 \\ &\quad + (r^x + ir^y)\,\boldsymbol{e}_1 \otimes \boldsymbol{\omega}^0 + (1 - r^z)\,\boldsymbol{e}_1 \otimes \boldsymbol{\omega}^1\Big) \end{aligned}$$

where $\boldsymbol{e}_i$ and $\boldsymbol{\omega}^i$ are basis vectors and forms in this new 2-dimensional complex vector space into which we have unpacked our quaternions. What they are will transpire when we have a closer look at some specific qubit states we know and understand well by now.

Consider first a state that is described by

$$\boldsymbol{r} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

or, in other words, by $r = e_z$, where e_z is a unit-length vector that points in the z direction, i.e., vertically up. This is a fully polarized, i.e., pure state. Its fiducial representation is $p = \frac{1}{2}(\varsigma_1 + \varsigma_z)$, its quaternion representation is $\rho = \frac{1}{2}(\mathbf{1} + \boldsymbol{\sigma}_z)$ and its density matrix representation is

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = e_0 \otimes \omega^0$$

This tells us that the pair e_0 and ω^0 can be employed to represent the $r = e_z$ state. We encountered this state in section 2.3 (page 52) that talked about polarized states and on that occasion we called it $|\uparrow\rangle$. We are going to adopt the same notation here including the complementary notation for the ω^0 form, namely:

$$\begin{aligned} e_0 &\doteq |\uparrow\rangle \\ \omega^0 &\doteq \langle\uparrow| \end{aligned}$$

e_0 and ω^0 represents the $r = e_z$ state.

Let us call the corresponding density operator $\rho_\uparrow$, to distinguish it from density operators that will describe other states, and we can write:

$$\rho_\uparrow = |\uparrow\rangle \otimes \langle\uparrow|$$

or

$$\rho_\uparrow = |\uparrow\rangle\langle\uparrow| \tag{4.10}$$

for short. Quantum physics researchers often drop the tensor product symbol $\otimes$[1].

$\rho_\uparrow = |\uparrow\rangle\langle\uparrow|$

Vector $|\uparrow\rangle$ and its matching form $\langle\uparrow|$ can be also described in terms of columns and rows of numbers, namely:

$$\begin{aligned} e_0 &\doteq |\uparrow\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \omega^0 &\doteq \langle\uparrow| \doteq (1,0) \\ e_0 \otimes \omega^0 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1,0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

The latter is a genuine matrix multiplication, i.e., multiplying a column vector by a row one, with the column vector on the left side of the row, and applying the usual matrix multiplication rules, builds up the 2×2 matrix that is in this case $\rho_\uparrow$. Also observe that

$\langle\uparrow|\uparrow\rangle = 1$

$$\langle\uparrow|\uparrow\rangle = \langle\omega^0, e_0\rangle = (1,0)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1$$

which is as it should be, because ω^0 is dual to e_0.

The terminology of *a form* and *a vector* adopted in this text is relatively

Forms, bras and rows versus vectors, kets and columns

---

[1]Some aren't even aware of its existence!

new to mathematics and newer still to physics (although it goes back a long way, at least to Cartan if not before). When Dirac invented his angular bracket notation used in quantum mechanics today, he called what we call vectors here *ket vectors* and he called what we call forms here *bra vectors*. This way a conjugation of a form and a vector, e.g., $\langle \uparrow | \downarrow \rangle$, becomes a *bra-ket*. Einstein and his differential geometry colleagues, on the other hand, called forms *covariant vectors* and vectors *contravariant vectors*, because form coefficients transform like basis vectors and vector coefficients transform the other way round. Finally, people who work with computers prefer to call forms *row vectors* and they call vectors *column vectors*. All these terminologies are still in use today, depending on who you talk to, and sometimes even depending on a context.

Now consider a state that is given by

$$\boldsymbol{r} = \left(\begin{array}{c} 0 \\ 0 \\ -1 \end{array} \right)$$

or, in other words, by $\boldsymbol{r} = -\boldsymbol{e}_z$. This is also a fully polarized state whose fiducial representation is $\boldsymbol{p} = \frac{1}{2}(\boldsymbol{\varsigma}_1 - \boldsymbol{\varsigma}_z)$ and whose quaternion representation is $\boldsymbol{\rho} = \frac{1}{2}(\mathbf{1} - \boldsymbol{\sigma}_z)$. Its density matrix representation is

$$\boldsymbol{\rho} = \left(\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \right) = \boldsymbol{e}_1 \otimes \boldsymbol{\omega}^1$$

This tells us that the pair $\boldsymbol{e}_1$ and $\boldsymbol{\omega}^1$ can be used to represent the $\boldsymbol{r} = -\boldsymbol{e}_z$ state. We encountered this state in section 2.3 too and called it $|\downarrow\rangle$ back then. So, we are going to adopt this notation here as well together with the complementary notation for its dual form:

$$\begin{array}{rcl} \boldsymbol{e}_1 & \doteq & |\downarrow\rangle \\ \boldsymbol{\omega}^1 & \doteq & \langle\downarrow| \end{array}$$

$\boldsymbol{\rho}_\downarrow = |\downarrow\rangle\langle\downarrow|$ Let us call the corresponding density operator $\boldsymbol{\rho}_\downarrow$. We can now write:

$$\boldsymbol{\rho}_\downarrow = |\downarrow\rangle \otimes \langle\downarrow| \quad \text{or simply} \quad \boldsymbol{\rho}_\downarrow = |\downarrow\rangle\langle\downarrow| \tag{4.11}$$

Vector $|\downarrow\rangle$ and its dual form $\langle\downarrow|$ can be described in terms of columns and rows of numbers as follows:

$$\begin{array}{rcl} \boldsymbol{e}_1 & \doteq & |\downarrow\rangle \doteq \left(\begin{array}{c} 0 \\ 1 \end{array} \right) \end{array}$$

$$\boldsymbol{\omega}_1 \;\dot{=}\; \langle\downarrow| \dot{=} (0,1)$$

$$\boldsymbol{e}_1 \otimes \boldsymbol{\omega}^1 \;=\; \begin{pmatrix} 0 \\ 1 \end{pmatrix}(0,1) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

where the latter, as before, is a genuine matrix multiplication. Since $|\downarrow\rangle$ and $\langle\downarrow|$ are dual, we have that

$$\langle\downarrow|\downarrow\rangle = 1$$

$\langle\downarrow|\downarrow\rangle = 1$

But we also have

$$\langle\downarrow|\uparrow\rangle = (0,1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

and

$$\langle\uparrow|\downarrow\rangle = (1,0) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$$

$\langle\downarrow|\uparrow\rangle = \langle\uparrow|\downarrow\rangle = 0$

And so we find that the two vectors, $|\uparrow\rangle$ and $|\downarrow\rangle$, comprise a basis of the 2-dimensional complex vector space in which $\boldsymbol{\rho}_\uparrow$ and $\boldsymbol{\rho}_\downarrow$ operate – once we have replaced the quaternion symbols $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ with Pauli matrices[2].

The vector space spanned by $|\downarrow\rangle$ and $|\uparrow\rangle$ is called the *Hilbert space* of a qubit. Apart from being a complex vector space and, in this case, 2-dimensional, it has some other properties that make it *Hilbert* and we're going to discover them one by one as we explore it.

Hilbert space of a qubit

Observe that vectors $|\uparrow\rangle$ and $|\downarrow\rangle$ also correspond to the *physical basis* states of the qubit, in the sense that was discussed in sections 1.5 (page 19) and 2.2 (page 46).

4.3   The superstition of superposition

What about qubit states such as $|\rightarrow\rangle$ and $|\otimes\rangle$, states that correspond to polarization directions that are perpendicular to $\boldsymbol{e}_z$? They are after all perfectly normal qubit beam states that can be *confirmed* by orienting the beam splitting magnet appropriately. Have they been excluded from the unitary formalism?

Since they can be described in terms of fiducial vectors, and therefore in terms of quaternions too, it should be possible to describe them within the framework of

---

[2] ... which are also called $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$. Whenever there is a possibility of confusion I will attempt to clarify whether the sigmas employed in various formulas should be thought of as quaternion symbols or matrices. The general rule is that if we operate on sigmas using their commutation and anticommutation properties only, they are quaternions. If we unpack them and use their matrix properties, they are Pauli matrices. Physicists call them Pauli matrices in all contexts.

the unitary formalism, even though the only basis vectors that we can build our vector space from correspond to $|\uparrow\rangle$ and $|\downarrow\rangle$.

Consider a density matrix that corresponds to $\boldsymbol{r} = \boldsymbol{e}_x$. The state is described by $\boldsymbol{p} = \frac{1}{2}\left(\boldsymbol{\varsigma}_1 + \boldsymbol{\varsigma}_x\right)$ or by $\boldsymbol{\rho} = \frac{1}{2}\left(\mathbf{1} + \boldsymbol{\sigma}_x\right)$. Unpon having unpacked $\boldsymbol{\sigma}_x$ we find:

$$
\begin{aligned}
\boldsymbol{\rho}_\rightarrow \ &= \ \frac{1}{2}\left(\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right) \\
&= \ \frac{1}{2}\left(|\uparrow\rangle \otimes \langle\uparrow| \ + \ |\uparrow\rangle \otimes \langle\downarrow| \ + \ |\downarrow\rangle \otimes \langle\uparrow| \ + \ |\downarrow\rangle \otimes \langle\downarrow|\right) \\
&= \ \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(\langle\uparrow| + \langle\downarrow|\right) \qquad (4.12)
\end{aligned}
$$

We can therefore write this very peculiar expression:

$$
|\rightarrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right)
$$

Similarly, for the qubit state that corresponds to $\boldsymbol{r} = -\boldsymbol{e}_x$ we get

$$
\begin{aligned}
\boldsymbol{\rho}_\leftarrow \ &= \ \frac{1}{2}\left(\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right) \\
&= \ \frac{1}{2}\left(|\uparrow\rangle \otimes \langle\uparrow| \ - \ |\uparrow\rangle \otimes \langle\downarrow| \ - \ |\downarrow\rangle \otimes \langle\uparrow| \ + \ |\downarrow\rangle \otimes \langle\downarrow|\right) \\
&= \ \frac{1}{\sqrt{2}}\left(|\uparrow\rangle - |\downarrow\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(\langle\uparrow| - \langle\downarrow|\right)
\end{aligned}
$$

Hence

$$
|\leftarrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle - |\downarrow\rangle\right)
$$

States $|\otimes\rangle$ and $|\odot\rangle$ that correspond do $\boldsymbol{r} = \boldsymbol{e}_y$ and $\boldsymbol{r} = -\boldsymbol{e}_y$ have similarly peculiar representation in the unitary space of a qubit but with one subtle difference. Let us start with $|\otimes\rangle = \frac{1}{2}\left(\boldsymbol{\varsigma}_1 + \boldsymbol{\varsigma}_y\right)$. The density matrix equivalent is:

$$
\begin{aligned}
\boldsymbol{\rho}_\otimes \ &= \ \frac{1}{2}\left(\begin{array}{cc} 1 & -i \\ i & 1 \end{array}\right) \\
&= \ \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + i\,|\downarrow\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(\langle\uparrow| - i\,\langle\downarrow|\right)
\end{aligned}
$$

At first glance it may look like we have a vector here and a form that are not related, or not *dual*. But observe this:

$$\frac{1}{\sqrt{2}}\left(\langle\uparrow| - i\langle\downarrow|\right)\frac{1}{\sqrt{2}}\left(|\uparrow\rangle + i\,|\downarrow\rangle\right)$$

$$= \frac{1}{2}\left(\langle\uparrow|\uparrow\rangle + i\,\langle\uparrow|\downarrow\rangle - i\,\langle\downarrow|\uparrow\rangle + \langle\downarrow|\downarrow\rangle\right)$$

$$= \frac{1}{2}\left(1 + i\,0 - i\,0 + 1\right) = 1$$

This suggests that in this particular 2-dimensional vector space with complex coefficients, whenever we want to make a form out of a vector, we need to convert all occurrences of i to $-i$. This is another feature of what's called a Hilbert space. If you want to be one, it's not enough to be just a complex vector space, you have to have this property too. Furthermore if we have an operator $\boldsymbol{A}$ acting on a vector $|\,\Psi\rangle$, namely $\boldsymbol{A}\,|\,\Psi\rangle$, then the image of this operation in the form world is $\langle\Psi\,|\,\boldsymbol{A}^{\dagger}$, where $\dagger$ represents *Hermitian conjugation* defined by equation (4.3) on page 112. Only if the operator is Hermitian, i.e., such that $\boldsymbol{A} = \boldsymbol{A}^{\dagger}$, we find that the dual of $\boldsymbol{A}\,|\,\Psi\rangle$ is $\langle\Psi\,|\,\boldsymbol{A}$ without the dagger.

In order to make a dual form in the Hilbert space we need to replace the vector's coefficients with their complex conjugates and then put the vector on its side.
The dual of $\boldsymbol{A}\,|\,\Psi\rangle$ is $\langle\Psi\,|\,\boldsymbol{A}^{\dagger}$

In summary:

$$|\otimes\rangle \;\; = \;\; \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + i\,|\downarrow\rangle\right) \quad\text{and}$$

$$\langle\otimes\,| \;\; = \;\; \frac{1}{\sqrt{2}}\left(\langle\uparrow| - i\langle\downarrow|\right)$$

And, without further calculation we can easily guess that

$$|\odot\rangle \;\; = \;\; \frac{1}{\sqrt{2}}\left(|\uparrow\rangle - i\,|\downarrow\rangle\right) \quad\text{and}$$

$$\langle\odot\,| \;\; = \;\; \frac{1}{\sqrt{2}}\left(\langle\uparrow| + i\langle\downarrow|\right)$$

What is the meaning of $|\rightarrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right)$?

First observe that the plus operator, $+$, used in this context is *not* the same plus we used in adding probability vectors or their corresponding quaternions. Back then adding two probability vectors of pure states (similarly for quaternions) would *always* result in a mixed state, unless the constituents were one and the same state. But here we add two pure states, which are not identical at all, and we end up with another pure state. This addition of two vectors in the qubit's Hilbert space, which is called a *superposition*, does *not* map onto addition of the two probability vectors, or quaternions. The transition

A plus in the superposition does not translate into a mixture of states

$$\boldsymbol{\rho}_{\rightarrow} \rightarrow |\rightarrow\rangle$$

is clearly non linear.

The second thing to ponder on is that this notational trick not so much reveals new physics, as hides some of what is transparent in the fiducial or quaternion formalisms, namely polarization states that are perpendicular to $\boldsymbol{e}_z$. The real and *complete* physics of a qubit in the $|\rightarrow\rangle$ state is described by

$$\boldsymbol{p}_{\rightarrow} = \frac{1}{2}\left(\boldsymbol{\varsigma}_1 + \boldsymbol{\varsigma}_x\right) = \begin{pmatrix} 1/2 \\ 1/2 \\ 1 \\ 1/2 \end{pmatrix}$$

This vector says that if a beam of qubits in this state is sent through a magnetic beam splitter with the magnetic field gradient pointing in the $\boldsymbol{e}_z$ direction, then the beam is going to split in half: one half of all incident qubits will swing upwards, the other half will swing downwards. If the beam splitter is rotated so that its magnetic field gradient points in the $\boldsymbol{e}_y$ direction, again the beam is going to split in two equal halves. But if the beam splitter is rotated so that its magnetic field gradient points in the $\boldsymbol{e}_x$ direction, then the beam is not going to split at all. Instead all qubits in the beam will be deflected in the same direction. This last measurement therefore *confirms* that the beam is polarized in the $\boldsymbol{e}_x$ direction.

"Unitarists" believe that a qubit that is in a superposition of two states is in both states simultaneously.
But people who get too obsessed with the notation of the unitary formalism have a different interpretation of this state. They will say that $|\rightarrow\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right)$ means that *every* qubit in the beam is *simultaneously* polarized in the $\boldsymbol{e}_z$ direction *and* in the $-\boldsymbol{e}_z$ direction – and they will argue that this is what the experiment shows: if you send a beam of qubits through the beam splitter oriented in the $\boldsymbol{e}_z$ direction, the beam splits in half – some qubits get deflected upwards, some get deflected downwards.

There is a huge leap of faith (rather than science) in their argument combined with emergency landing in thorny bushes of a religious dogma. First, they attribute to every individual qubit what is clearly a property *not* of an individual qubit but of a statistical ensemble of qubits. This is what every theory of probability is all about and quantum mechanics of a qubit is merely a yet another theory of probability. There is a certain probability that I may find a \$100 bill on the floor of the supermarket tomorrow. Does this mean that I half-have and half-not-have this bill in my hand today? Does this mean that I have a certain proclivity towards finding \$100 bills on supermarket floors?

Second, they tend to forget about the option of rotating the beam splitter so as to find a direction for which the beam no longer splits. This is because this

possibility cannot be clearly and easily read from the unitary representation of the state. It is there, but it's hidden.

On the other hand, the interpretation that is read from the fiducial vector $\boldsymbol{p}$ or from its equivalent quaternion $\boldsymbol{\rho}$ is that the qubits in the beam are *neither* polarized in the $\boldsymbol{e}_z$ direction, *nor* are they polarized in the $-\boldsymbol{e}_z$ direction. They are polarized in the $\boldsymbol{e}_x$ direction instead, which is perpendicular to both $\boldsymbol{e}_z$ and $-\boldsymbol{e}_z$. They are going to be flipped onto either the $\boldsymbol{e}_z$ or the $-\boldsymbol{e}_z$ directions only when they are measured with such probability as can be read from $\boldsymbol{p}$ or $\boldsymbol{\rho}$.

"Fiducialists" believe that a qubit that is in a superposition of two states is in neither of the states, instead it is in a third state that is altogether different.

It is not always easy to find a clear, physical interpretation of a quantum state that is described by a linear combination of some basis vectors in the Hilbert space. For example, what does $|\rightarrow\rangle$ mean for the quantronium? But rather than saying that the quantum object has *all* the constituent properties of a superposition, it is probably better to say that it has *neither* of them, that its property is altogether different. What it is exactly we're not going to see until we switch back to the density operator formalism (e.g., through the $|\Psi\rangle \rightarrow |\Psi\rangle\langle\Psi|$ operation) or even better to the fiducial formalism (throuth the $\boldsymbol{\sigma}_i \rightarrow \boldsymbol{\varsigma}_i$ operation) and consider probabilities of every experimental measurement needed to fully characterize the state. It may sometimes happen that an investigated system, e.g., a quantronium, is not equipped to carry out a full set of measurements and all it can deliver are p^0 and p^1, but, at least in principle, p^2 and p^3 are there too and they can be detected with appropriately improved measurement setup, or by more elaborate experimentation. For example, Ramsey fringes generated by the quantronium demonstrated to us the presence of Larmor precession, even though the quantronium does not have a circuitry needed to measure p^2 and p^3.

> N dimensional quantum systems, where N is the number of dimensions of their corresponding Hilbert space map onto N^2 dimensional fiducial systems [26]. Hence, the number of parameters needed to fully characterize a quantum system grows very rapidly with the dimension of the system. This is a practical reason why we often prefer to work within the confines of the unitary formalism – especially for more complex systems. Whereas it is easy to talk about p^0, p^1, p^2 and p^3 for a single qubit, if we were to consider an 8-qubit register, the number of dimensions of the corresponding Hilbert space would be $2^8 = 256$, but the number of dimensions of the corresponding fiducial space would be 65,536. Consequently, even if the "Fiducialists" are right, the "Unitarists" are more practical.

Consider a general superposition of the two Hilbert space basis vectors $|\uparrow\rangle$ and

Fiducial vector of a general state in the qubit's Hilbert space

$|\downarrow\rangle$:

$$| \, x \rangle = a \, |\uparrow\rangle + b \, |\downarrow\rangle$$

where a and b are two complex numbers. We can convert this vector to its corresponding density matrix by constructing the form

$$\langle x \, | = a^* \, \langle\uparrow| + b^* \, \langle\downarrow|$$

where the asterisk denotes complex conjugation, i.e., $i \to -i$, and then by building an operator out of the two:

$$| \, x \rangle\langle x \, | = aa^* \, |\uparrow\rangle\langle\uparrow| + ab^* \, |\uparrow\rangle\langle\downarrow| + ba^* \, |\downarrow\rangle\langle\uparrow| + bb^* \, |\downarrow\rangle\langle\downarrow|$$

Let us compare this to the general form of the density matrix as obtained from the quaternion representation in order to *find* how the full information about the state of the polarization of the qubit beam is hidden in the unitary formalism's superposition:

$$
\begin{aligned}
aa^* &= \frac{1}{2}\left(1 + r^z\right) \\
bb^* &= \frac{1}{2}\left(1 - r^z\right) \\
ab^* &= \frac{1}{2}\left(r^x - ir^y\right) \\
ba^* &= \frac{1}{2}\left(r^x + ir^y\right)
\end{aligned}
$$

Adding the first two equations yields an interesting condition:

$$aa^* + bb^* = 1$$

This tells us that we cannot construct just any linear combinations of $|\uparrow\rangle$ and $|\downarrow\rangle$. The combinations that are physically meaningful are restricted to such that *Physically meaningful Hilbert space vectors must be normalized* $aa^* + bb^* = 1$. This is called a normalization condition. All Hilbert space vectors that correspond to physical states of a qubit must satisfy it.

Also observe that $aa^* = \frac{1}{2}\left(1 + r^z\right) = p^0$ is the probability of registering the qubit with its spin up and $bb^* = \frac{1}{2}\left(1 - r^z\right) = p^1$ is the probability of registering the qubit with its spin down. The normalization condition is therefore the same as $p^0 + p^1 = 1$.

Subtracting the second equation from the first one yields

$$aa^* - bb^* = r^z \tag{4.13}$$

We can extract r^x and r^y similarly from the third and the fourth equations:

$$ab^* + ba^* = r^x \quad \text{and} \tag{4.14}$$

$$i\left(ab^* - ba^*\right) = r^y \tag{4.15}$$

Now let us see if we can fill the whole Bloch *ball* with superpositions of $|\uparrow\rangle$ and $|\downarrow\rangle$:

Only the Bloch sphere, not the full ball, can be covered with Hilbert space states

$$
\begin{aligned}
&r^x r^x + r^y r^y + r^z r^z \\
&= (ab^* + ba^*)^2 + i^2 \left(ab^* - ba^*\right)^2 + (aa^* - bb^*)^2 \\
&= (ab^*)^2 + 2ab^*ba^* + (ba^*)^2 \\
&\quad - (ab^*)^2 + 2ab^*ba^* - (ba^*)^2 \\
&\quad + (aa^*)^2 - 2aa^*bb^* + (bb^*)^2 \\
&= (aa^*)^2 + 2aa^*bb^* + (bb^*)^2 \\
&= (aa^* + bb^*)^2 \\
&= 1
\end{aligned}
$$

We can fill the Bloch sphere, but *not* the Bloch ball. In other words, the unitary formalism is restricted to fully polarized beams only. We have no means to describe mixtures within the framework of this formalism. The unitary formalism represents only a subset of the quantum probability theory and as such it seems incomplete.

Mixtures of quantum states cannot be described within the confines of the unitary formalism

This would be a high price to pay. If we cannot describe mixtures, we cannot describe the process of the measurement, since, as we had seen in chapter 2, section 2.5 (page 59), the measurement process converted a pure state to a mixture.

Within the framework of the unitary formalism the measurement process has been handled traditionally by introducing an auxiliary unphysical *axiom*. The axiom is unphysical, because it elevates the measurement process above the theory making it a special act, rather than a result of a physical interaction between a measuring apparatus and an observed quantum system.

Measurement in the unitary formalism

A similar problem would seem to affect our ability to describe the effects of the interaction between an observed quantum system and the environment within the framework of unitary formalism. We saw in section 2.11 (page 83) that talked about the quantronium circuit that the qubit depolarized gradually. This manifested in the diminishing amplitude of Rabi and Ramsey oscillations. The initially pure state of the qubit converted into a mixture.

Depolarization in the unitary formalism

But things are not so desperate and the unitary formalism is more useful than it seems at first glance[3].

The situation here is somewhat similar to the situation we encounter in investigating classical dissipative systems. There energy is obviously lost to the environment, yet it does not mean that energy on the whole is not conserved. The energy may leak from an observed subsystem, e.g., a damped oscillator, but we can recover it eventually by including air and its expansion, friction, heat, as well as temperature and lengthening of the spring in the model.

Dissipative quantum systems can be embedded in larger systems that are unitary.

It is similarly with the unitary formalism. We will show in the next chapter that a "dissipative", i.e., a non-unitary quantum system can be always embedded in a larger "non-dissipative" unitary system. This way, unitarity, like energy, is "conserved" globally, even if it appears to leak out of the portion of the system under observation. Careful analysis of what happens when a unitary quantum system interacts with the environment lets us derive dissipative quantum equations such as the Lindblad equation – its classical analog would be a Newton equation with friction.

The universe as a unitary quantum system

This mathematical trick leads some physicists to proclaim that the universe itself must be a quantum unitary system, but this is just a religious belief with no grounding in observations or laboratory experiments. On the contrary, recent observations seem to suggest that gravity somehow leaks out of the universe (and there is no convincing quantum theory of gravity to begin with), which manifests in its accelerating expansion. If gravity can leak out, then why not energy and why not unitarity? Although we know a great deal about the universe, there is apparently much that we don't know either.

Further specification of Pauli matrices

Let us go back to the choice of $x = 1$ in equations (4.1) and (4.2) in section 4.1 (page 111) and explain it some more.

Consider again equation (4.12) on page 118, but this time let us use

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & x \\ x^{-1} & 0 \end{pmatrix}$$

This would yield:

$$\begin{aligned} \boldsymbol{\sigma}_\rightarrow &= \frac{1}{2}\begin{pmatrix} 1 & x \\ x^{-1} & 1 \end{pmatrix} \\ &= \frac{1}{2}\left(|\uparrow\rangle \otimes \langle\uparrow| + x\,|\uparrow\rangle \otimes \langle\downarrow| + x^{-1}\,|\downarrow\rangle \otimes \langle\uparrow| + |\downarrow\rangle \otimes \langle\downarrow|\right) \end{aligned}$$

---

[3]Not to mention the minor fact that nearly everything that's been done in quantum mechanics in the last 80 years or so was done with the unitary formalism.

To find a Hilbert space vector and its dual form that correspond to this operator we can try to match it against a tensor product of a general superposition and its dual form:

$$(a \ |\uparrow\rangle \ + \ b \ |\downarrow\rangle) \otimes (a^* \ \langle\uparrow| \ + b^* \ \langle\downarrow|)$$
$$= aa^* \ |\uparrow\rangle \otimes \langle\uparrow| \ + ab^* \ |\uparrow\rangle \otimes \langle\downarrow| \ + ba^* \ |\downarrow\rangle \otimes \langle\uparrow| \ + bb^* \ |\downarrow\rangle \otimes \langle\downarrow|$$

This yields the following equations:

$$aa^* \ = \ |a|^2 = \frac{1}{2}, \quad \text{hence} \quad |a| = \frac{1}{\sqrt{2}}$$

$$bb^* \ = \ |b|^2 = \frac{1}{2}, \quad \text{hence} \quad |b| = \frac{1}{\sqrt{2}}$$

$$ab^* \ = \ \frac{1}{2}x \tag{4.16}$$

$$ba^* \ = \ \frac{1}{2}x^{-1} \tag{4.17}$$

Let us use Euler notation for a and b, i.e.,

$$a \ = \ |a|e^{i\phi_a} \quad \text{and}$$
$$b \ = \ |b|e^{i\phi_b}$$

then equations (4.16) and (4.17) become

$$ab^* \ = \ |a||b|e^{i(\phi_a - \phi_b)} = \frac{1}{2}e^{i(\phi_a - \phi_b)} = \frac{1}{2}x$$
$$ba^* \ = \ |b||a|e^{i(\phi_b - \phi_a)} = \frac{1}{2}e^{-i(\phi_a - \phi_b)} = \frac{1}{2}x^{-1}$$

which yields

$$x = e^{i(\phi_a - \phi_b)}$$

If we were to restrict ourselves to a real x, the only choice for us would be $x = \pm 1$ and this would land us exactly where we are already with the only freedom left as to the sign in front of the $\boldsymbol{\sigma}_x$ matrix. We could also make a purely imaginary choice $x = \pm i$, and this would merely swap $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_y$. A non-trivial choice as to x is possible too. In this case x would have to be a complex number of length 1 given by

$$x = e^{i\phi_x} \tag{4.18}$$

Any other choice of x would make it impossible for us to recover the $|\rightarrow\rangle$ state within the resulting formalism, i.e., we could not identify a single Hilbert space vector that would correspond to it.

Observe also that choosing $x = e^{i\phi_x}$, which is the only choice we're ultimately left with, results in a Hermitian representation of $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_y$ namely:

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & e^{i\phi_x} \\ e^{-i\phi_x} & 0 \end{pmatrix}$$

$$\boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -ie^{i\phi_x} \\ ie^{-i\phi_x} & 0 \end{pmatrix} = \begin{pmatrix} 0 & e^{i(\phi_x - \pi/2)} \\ e^{-i(\phi_x - \pi/2)} & 0 \end{pmatrix}$$

If we were to choose $x = e^{i\phi_x}$ with $\phi_x \neq 0$, the coefficients a and b in $|\!\rightarrow\rangle = a\,|\!\uparrow\rangle + b\,|\!\downarrow\rangle$ would have to be such that $|a| = |b| = 1/\sqrt{2}$ and

$$\phi_a - \phi_b = \phi_x \tag{4.19}$$

Physical states in the Hilbert space are defined up to a constant phase factor.

In particular, observe that for $\phi_x = 0$, which is equivalent to our choice of $x = 1$, we end up with $\phi_a = \phi_b$, but we are not forced to take $\phi_a = \phi_b = 0$. This means that we can multiply the superposition by $e^{i\phi_a} = e^{i\phi_b} = e^{i\phi}$ and this is still going to yield the same physical state.

4.4   Probability amplitudes

A probability that a qubit that is in a *pure* state defined by $\boldsymbol{r}_1$, where $\boldsymbol{r}_1 \cdot \boldsymbol{r}_1 = 1$, is going to be filtered onto an "up" beam in some measuring apparatus is given by $p^0 = (1 + r_1^z)/2$. What is a probability that the qubit is going to be filtered onto an "up" beam in a differently oriented apparatus? Suppose that the orientation of the apparatus is $\boldsymbol{r}_2$, where $\boldsymbol{r}_2 \cdot \boldsymbol{r}_2 = 1$. We can answer this question by rotating our system of coordinates so that $\boldsymbol{r}_2 = \boldsymbol{e}_{z'}$ and now we have that $p^{0'} = \left(1 + r_1^{z'}\right)/2$. But what is $r_1^{z'}$? Recall that $r^z = \langle \boldsymbol{\omega}^z, \boldsymbol{r} \rangle = \boldsymbol{e}_z \cdot \boldsymbol{r}$. This holds for every orthonormal basis $\boldsymbol{e}_i$, including the new basis defined by $\boldsymbol{r}_2$. Consequently $r_1^{z'} = \boldsymbol{e}_{z'} \cdot \boldsymbol{r}_1 = \boldsymbol{r}_2 \cdot \boldsymbol{r}_1$. The probability of a *transition* from a *pure* state that corresponds to $\boldsymbol{r}_1$ to a *pure* state that corresponds to $\boldsymbol{r}_2$ is therefore:

Probability of a transition from $\boldsymbol{r}_1$ to $\boldsymbol{r}_2$

$$p_{\boldsymbol{r}_2 \leftarrow \boldsymbol{r}_1} = \frac{1}{2}\left(1 + \boldsymbol{r}_2 \cdot \boldsymbol{r}_1\right)$$

Let us digress here for a moment. The above formula is expressed in terms of fiducial vector parametrizations. Can we express it instead in terms of fiducial vectors themselves?

Let the two states be described by

$$\boldsymbol{p}_1 = \frac{1}{2}\left(\boldsymbol{\varsigma}_1 + r_1^x\boldsymbol{\varsigma}_x + r_1^y\boldsymbol{\varsigma}_y + r_1^z\boldsymbol{\varsigma}_z\right) \quad \text{and}$$

$$\boldsymbol{p}_2 = \frac{1}{2}\left(\boldsymbol{\varsigma}_1 + r_2^x\boldsymbol{\varsigma}_x + r_2^y\boldsymbol{\varsigma}_y + r_2^z\boldsymbol{\varsigma}_z\right)$$

Consider $\langle \tilde{\boldsymbol{p}}_2, \boldsymbol{p}_1 \rangle$

$$
\begin{aligned}
\langle \tilde{\boldsymbol{p}}_2, \boldsymbol{p}_1 \rangle &= \frac{1}{4} \left\langle \boldsymbol{\varsigma}^1 + r_2^x \boldsymbol{\varsigma}^x + r_2^y \boldsymbol{\varsigma}^y + r_2^z \boldsymbol{\varsigma}^z, \boldsymbol{\varsigma}_1 + r_1^x \boldsymbol{\varsigma}_x + r_1^y \boldsymbol{\varsigma}_y + r_1^z \boldsymbol{\varsigma}_z \right\rangle \\
&= \frac{1}{4} \left(\langle \boldsymbol{\varsigma}^1, \boldsymbol{\varsigma}_1 \rangle + r_1^x r_2^x \langle \boldsymbol{\varsigma}^x, \boldsymbol{\varsigma}_x \rangle + r_1^y r_2^y \langle \boldsymbol{\varsigma}^y, \boldsymbol{\varsigma}_y \rangle + r_1^z r_2^z \langle \boldsymbol{\varsigma}^z, \boldsymbol{\varsigma}_z \rangle \right) \\
&= \frac{1}{2} \left(1 + \boldsymbol{r}_1 \cdot \boldsymbol{r}_2 \right)
\end{aligned}
$$

because $\langle \boldsymbol{\varsigma}^i, \boldsymbol{\varsigma}_j \rangle = 2\delta^i{}_j$. We will soon see that there is a very nice typographic correspondence between

$$
p_{2 \leftarrow 1} = \langle \tilde{\boldsymbol{p}}_2, \boldsymbol{p}_1 \rangle \tag{4.20}
$$

and its unitary equivalent.

Although we have derived this formula for *pure* states only, it may be extended to a situation in which $\boldsymbol{r}_1$ is *any* state, possibly a mixed one, and $\boldsymbol{r}_2$ is pure. The latter is required, because in our derivation we really thought of $\boldsymbol{r}_2$ as defining a direction and we made use of the fact that its length is 1 (when we equated $\boldsymbol{r}_2$ and $\boldsymbol{e}_{z'}$). But we never made any use of $\boldsymbol{r}_1$ being of length 1.

Let us substitute the unitary description coefficients a_1, b_1, a_2 and b_2 in place of r_1^x, r_1^y, r_1^z, r_2^x, r_2^y and r_2^z. Using equations (4.13), (4.14) and (4.15) yields

$$
\begin{aligned}
\boldsymbol{r}_2 \cdot \boldsymbol{r}_1 = \ & (a_2 b_2^* + b_2 a_2^*)(a_1 b_1^* + b_1 a_1^*) \\
& - (a_2 b_2^* - b_2 a_2^*)(a_1 b_1^* - b_1 a_1^*) \\
& + (a_2 a_2^* - b_2 b_2^*)(a_1 a_1^* - b_1 b_1^*)
\end{aligned} \tag{4.21}
$$

Also recall that for every fully polarized state

$$
aa^* + bb^* = 1
$$

So we can replace the 1 in $(1 + \boldsymbol{r}_2 \cdot \boldsymbol{r}_1)$ with

$$
1 = 1 \cdot 1 = (a_2 a_2^* + b_2 b_2^*)(a_1 a_1^* + b_1 b_1^*) \tag{4.22}
$$

Let us now combine (4.21) and (4.22):

$$
\begin{aligned}
1 + \boldsymbol{r}_2 \cdot \boldsymbol{r}_1 \\
= \ & (a_2 a_2^* + b_2 b_2^*)(a_1 a_1^* + b_1 b_1^*) \\
& + (a_2 b_2^* + b_2 a_2^*)(a_1 b_1^* + b_1 a_1^*) \\
& - (a_2 b_2^* - b_2 a_2^*)(a_1 b_1^* - b_1 a_1^*) \\
& + (a_2 a_2^* - b_2 b_2^*)(a_1 a_1^* - b_1 b_1^*)
\end{aligned}
$$

Let us have a sanguine look at this equation. Observe that the first line is much like the last one, but there are minuses in front of the b-terms in the last one. Also, the third line is much like the second one, but again there are minuses in front of the ba-terms. On the other hand there is a minus in front of the third line, so we should really flip the minuses and the pluses inside it. All these minuses will result in four merry cancellations, so that the final result will be left with 4 terms only. And it's easy to see what they're going to be: The middle terms that result from the full expansion of the first and the last line will drop out, and the edge terms that result from the full expansion of the second and the third line will drop out too, leaving us with

$$1 + \boldsymbol{r}_2 \cdot \boldsymbol{r}_1 = 2a_2 a_2^* a_1 a_1^* + 2b_2 b_2^* b_1 b_1^* + 2a_2 b_2^* b_1 a_1^* + 2b_2 a_2^* a_1 b_1^* \tag{4.23}$$

The $1/2$ in front of $\frac{1}{2}\left(1 + \boldsymbol{r}_2 \cdot \boldsymbol{r}_1\right)$ will kill the 2-s in (4.23), and so we're left with

$$\begin{aligned} p_{\boldsymbol{r}_2 \leftarrow \boldsymbol{r}_1} &= a_2 a_2^* a_1 a_1^* + b_2 b_2^* b_1 b_1^* + a_2 b_2^* b_1 a_1^* + b_2 a_2^* a_1 b_1^* \\ &= \left(a_2^* a_1 + b_2^* b_1\right)\left(a_2 a_1^* + b_2 b_1^*\right) \end{aligned}$$

First, observe that the second component in this product is a complex conjugate of the first component. Also observe that the first component is

$$a_2^* a_1 + b_2^* b_1 = \left\langle a_2^* \langle \uparrow | + b_2^* \langle \downarrow | \, \middle| \, a_1 | \uparrow \rangle + b_1 | \downarrow \rangle \right\rangle = \langle \Psi_2 \mid \Psi_1 \rangle$$

and the second component is

$$a_2 a_1^* + b_2 b_1^* = \left\langle a_1^* \langle \uparrow | + b_1^* \langle \downarrow | \, \middle| \, a_2 | \uparrow \rangle + b_2 | \downarrow \rangle \right\rangle = \langle \Psi_1 \mid \Psi_2 \rangle$$

$\langle \Psi_2 \mid \Psi_1 \rangle$ *is the probability amplitude of transition from* $\boldsymbol{r}_1$ *to* $\boldsymbol{r}_2$.

And so, we have discovered that:

- $\langle \Psi_2 \mid \Psi_1 \rangle = \langle \Psi_1 \mid \Psi_2 \rangle^*$ – this should not come as a surprise, because we have already discovered that when converting a vector into its dual form we needed to replace the vector coefficients with their complex conjugates. This result is merely a consequence of this.

-
$$p_{\boldsymbol{r}_2 \leftarrow \boldsymbol{r}_1} = \langle \tilde{\boldsymbol{p}}_2, \boldsymbol{p}_1 \rangle = |\langle \Psi_2 \mid \Psi_1 \rangle|^2 \tag{4.24}$$

For this reason the complex number $\langle \Psi_2 \mid \Psi_1 \rangle$ is called a *probability amplitude* for the transition from $| \Psi_1 \rangle$ to $| \Psi_2 \rangle$. The tradition in quantum mechanics, that goes back all the way to Dirac, is to write and read probability amplitude expressions from right to left.

A special case here are probability amplitudes for transitions to the basis states $\langle\uparrow|\ \Psi_1\rangle$ and $\langle\downarrow|\ \Psi_1\rangle$. These amplitudes are simply:

$$\langle\uparrow|\ \Psi_1\rangle \quad = \quad \left\langle\langle\uparrow|\ \Big|\ a_1\ |\uparrow\rangle + b_1\ |\downarrow\rangle\right\rangle = a_1 \qquad (4.25)$$

$$\langle\downarrow|\ \Psi_1\rangle \quad = \quad \left\langle\langle\downarrow|\ \Big|\ a_1\ |\uparrow\rangle + b_1\ |\downarrow\rangle\right\rangle = b_1 \qquad (4.26)$$

We can therefore write

$$\begin{aligned}
|\ \Psi_1\rangle &= |\uparrow\rangle\langle\uparrow|\ \Psi_1\rangle + |\downarrow\rangle\langle\downarrow|\ \Psi_1\rangle \\
|\ \Psi_2\rangle &= |\uparrow\rangle\langle\uparrow|\ \Psi_2\rangle + |\downarrow\rangle\langle\downarrow|\ \Psi_2\rangle \\
\langle\Psi_2\ |\ \Psi_1\rangle &= \langle\Psi_2\ |\uparrow\rangle\langle\uparrow|\ \Psi_1\rangle + \langle\Psi_2\ |\downarrow\rangle\langle\downarrow|\ \Psi_1\rangle
\end{aligned}$$

This last expression has a peculiar interpretation within the lore of quantum mechanics that derives from the "superstition of superposition". It says that on its way from $|\ \Psi_1\rangle$ to $|\ \Psi_2\rangle$ a qubit transits both through $|\uparrow\rangle$ and $|\downarrow\rangle$ at the same time, as if splitting itself and being in these two states simultaneously. Well, as we have seen before, a qubit that is described by $|\ \Psi_1\rangle = a_1\ |\uparrow\rangle + b_1\ |\downarrow\rangle$ is *neither* in the $|\uparrow\rangle$ *More* state *nor* in the $|\downarrow\rangle$ state. It is in a different state altogether that is not clearly ex- *superstition* pressed by the unitary formalism, but, instead, it is hidden inside the two complex numbers a_1 and b_1. This state can be always extracted by switching to the density operator formalism and extracting probabilities of all required qubit characteristics from it, or by juggling a_1 and b_1 following equations (4.13), (4.14) and (4.15).

The word "transition" hints at some "motion" with intermediate states in the system. But when we think about a state of a qubit as polarization and then the act of the measurement as filtration, there isn't really any motion involved. Some qubits get through the filter, some don't. This is the same as filtering photons by a polarizer plate. What is the actual mechanism involved in getting through the filter and what it is exactly that happens to a qubit when it gets rejected, this is something that quantum mechanics doesn't really tell us much about. It is a probability theory and as such it describes rather than explains[4].

When contemplating quantum transitions, it is probably best to think of $\langle\Psi_2\ |\ \Psi_1\rangle$ as representing just such an act of filtration. The qubit does not really go through $|\uparrow\rangle$ and $|\downarrow\rangle$ any more than it exists in the $|\uparrow\rangle$ and the $|\downarrow\rangle$ states at the same time. These are just mathematical expressions written on paper that derive from

---

[4]This can be said about any other physics theory. Even though some of them may seem to explain things, it is enough that a new, more accurate theory is discovered to relagate the old one to a mere phenomenology. At the end of the day it is safer to leave the difference between "describe" and "explain" to philosophers and focus on "predict" instead.

the linearity of the theory. It is perhaps better to think of the qubit as ceasing to be in the $\mid \Psi_1 \rangle$ state and re-appearing in the $\mid \Psi_2 \rangle$ state without anything in-between. Its quantum mechanical "transition" is not continuous.

But one should not confuse transition probability amplitudes with a unitary evolution of a qubit. The unitary Hamiltonian evolution is continuous. There are no sudden jumps here. The transition amplitudes we talk about refer to probabilities of registering the qubit in such or another state. They refer to the act of measurement, the act that in itself is not unitary.

4.5   Spinors

How do qubit representations in the Hilbert space transform under the change of the canonical basis in the qubit's physical space?

To answer this question we must first answer another question. How do the probabilities encapsulated in the qubit's fiducial vector change in the same context? The fiducial vector of a qubit is

$$
\boldsymbol{p} = \frac{1}{2} \begin{pmatrix} 1 + r^z \\ 1 - r^z \\ 1 + r^x \\ 1 + r^y \end{pmatrix}
$$

The components r^x, r^y and r^z depend on the specific choice of a basis in the qubit's 3D space. This 3D space may be a geometric space, as is the case for, say, neutron spin, or it may be some other more abstract space, as is the case for the quantronium circuit. But mathematically they're all the same: 3D real vector spaces.

We can rewrite $\boldsymbol{p}$ as follows

$$
\boldsymbol{p} = \frac{1}{2} \begin{pmatrix} 1 + \langle \boldsymbol{\omega}^z, \boldsymbol{r} \rangle \\ 1 - \langle \boldsymbol{\omega}^z, \boldsymbol{r} \rangle \\ 1 + \langle \boldsymbol{\omega}^x, \boldsymbol{r} \rangle \\ 1 + \langle \boldsymbol{\omega}^y, \boldsymbol{r} \rangle \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \boldsymbol{e}_z \cdot \boldsymbol{r} \\ 1 - \boldsymbol{e}_z \cdot \boldsymbol{r} \\ 1 + \boldsymbol{e}_x \cdot \boldsymbol{r} \\ 1 + \boldsymbol{e}_y \cdot \boldsymbol{r} \end{pmatrix}
$$

where $\boldsymbol{\omega}^i$ and $\boldsymbol{e}_i$, $i = x, y, z$ are canonical basis forms and vectors, such that $\boldsymbol{e}_z$ points in the direction of the "magnetic field" used to measure the qubit.

If we rotate the canonical basis so that $\boldsymbol{e}_i \to \boldsymbol{e}_{i'}$ then $\boldsymbol{p}$ is going to change, but $\boldsymbol{r}$ will remain the same, because it represents the qubit and its physics, neither of which should depend on our choice of directions. Hence

$$\boldsymbol{p}' = \frac{1}{2} \begin{pmatrix} 1 + \boldsymbol{e}_{z'} \cdot \boldsymbol{r} \\ 1 - \boldsymbol{e}_{z'} \cdot \boldsymbol{r} \\ 1 + \boldsymbol{e}_{x'} \cdot \boldsymbol{r} \\ 1 + \boldsymbol{e}_{y'} \cdot \boldsymbol{r} \end{pmatrix}$$

represents the qubit in the same state $\boldsymbol{r}$. The only thing that would have changed is the way we look at it – through a different canonical basis $\boldsymbol{e}_{i'}$ in the qubit's physical space.

Suppose we rotate the basis by θ in the $\boldsymbol{e}_x \times \boldsymbol{e}_z$ plane as shown in figure 4.1. The basis vectors $\boldsymbol{e}_x$, $\boldsymbol{e}_z$, $\boldsymbol{e}_{x'}$ and $\boldsymbol{e}_{z'}$ all have length 1. In this operation vector
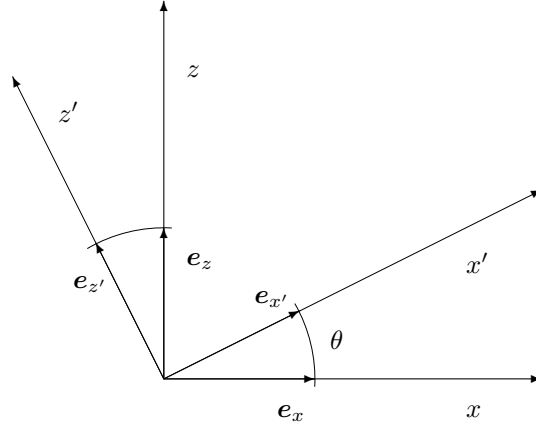
Figure 4.1: A counter-clockwise rotation of basis $\boldsymbol{e}_i$ onto $\boldsymbol{e}_{i'}$ by angle θ.

$\boldsymbol{e}_y$ remains unchanged, i.e., $\boldsymbol{e}_y = \boldsymbol{e}_{y'}$. The new basis vectors $\boldsymbol{e}_{x'}$ and $\boldsymbol{e}_{z'}$ can be expressed in terms of $\boldsymbol{e}_x$ and $\boldsymbol{e}_z$, namely

$$\boldsymbol{e}_{x'} = \cos\theta\,\boldsymbol{e}_x + \sin\theta\,\boldsymbol{e}_z \tag{4.27}$$

$$\boldsymbol{e}_{z'} = -\sin\theta\,\boldsymbol{e}_x + \cos\theta\,\boldsymbol{e}_z \tag{4.28}$$

The probability vector $\boldsymbol{p}$ is therefore going to change as follows

$$\boldsymbol{p}' = \frac{1}{2} \begin{pmatrix} 1 + \boldsymbol{e}_{z'} \cdot \boldsymbol{r} \\ 1 - \boldsymbol{e}_{z'} \cdot \boldsymbol{r} \\ 1 + \boldsymbol{e}_{x'} \cdot \boldsymbol{r} \\ 1 + \boldsymbol{e}_{y'} \cdot \boldsymbol{r} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 - \sin\theta\, r^x + \cos\theta\, r^z \\ 1 + \sin\theta\, r^x - \cos\theta\, r^z \\ 1 + \cos\theta\, r^x + \sin\theta\, r^z \\ 1 + r^y \end{pmatrix} \tag{4.29}$$

Let us focus on a very simple case. Suppose that $\boldsymbol{r} = \boldsymbol{e}_z$ (i.e., $r^x = r^y = 0$). The unitary representation of this state is $|\uparrow\rangle$. When looked at from the new basis $\boldsymbol{e}_{i'}$ the probability vector evaluates to:

$$\boldsymbol{p}' = \frac{1}{2} \begin{pmatrix} 1 + \cos\theta \\ 1 - \cos\theta \\ 1 + \sin\theta \\ 1 \end{pmatrix}$$

In the unitary representation that corresponds to the primed basis, let's call it $|\uparrow\rangle'$ and $|\downarrow\rangle'$, the state is going to be described by

$$|\uparrow\rangle = a\,|\uparrow\rangle' + b\,|\downarrow\rangle'$$

where a and b must satisfy equations (4.13), (4.14) and (4.15) from page 122, as well as the normalization condition. In this case

$$
\begin{aligned}
aa^* + bb^* &= 1 \\
aa^* - bb^* &= r^{z'} = \cos\theta \\
ab^* + ba^* &= r^{x'} = \sin\theta \\
i(ab^* - ba^*) &= r^{y'} = 0
\end{aligned}
$$

The last condition tells us that ab^* is real. This implies that a and b share the same phase angle. But since Hilbert space vectors are defined up to a constant phase factor anyway, we can just as well ignore this phase factor and assume that both a and b are real. This greatly simplifies our algebra:

$$
\begin{aligned}
a^2 + b^2 &= 1 \\
a^2 - b^2 &= r^{z'} = \cos\theta \\
2ab &= r^{x'} = \sin\theta
\end{aligned}
$$

From the normalization condition

$$b^2 = 1 - a^2$$

hence

$$a^2 - b^2 = 2a^2 - 1 = \cos\theta$$

hence

$$a^2 = \frac{\cos\theta + 1}{2} = \cos^2\frac{\theta}{2}$$

and

$$b^2 = 1 - a^2 = 1 - \cos^2\frac{\theta}{2} = \sin^2\frac{\theta}{2}$$

In summary:

$$a = \pm\cos\frac{\theta}{2} \quad\text{and}\quad b = \pm\sin\frac{\theta}{2}$$

How to choose the signs? For $\theta = 0$ we must have

$$|\uparrow\rangle = |\uparrow\rangle'$$

hence $a = +\cos\frac{\theta}{2}$. Then the $r^{x'}$ equation tells us that for small positive angles θ b must have the same sign as a, consequently $b = +\sin\frac{\theta}{2}$.

In summary

$$|\uparrow\rangle = \cos\frac{\theta}{2}\ |\uparrow\rangle' + \sin\frac{\theta}{2}\ |\downarrow\rangle' \tag{4.30}$$

If $\boldsymbol{r} = -\boldsymbol{e}_z$, then the reasoning is very similar, but we end up with minuses in front of cos and sin in the $r^{z'}$ and $r^{x'}$ equations. The latter tells us that this time a and b must be of the opposite sign, and the former selects sin for a, and so, expecting that for $\theta = 0$ we should have $|\downarrow\rangle = |\downarrow\rangle'$ yields

$$|\downarrow\rangle = -\sin\frac{\theta}{2}\ |\uparrow\rangle' + \cos\frac{\theta}{2}\ |\downarrow\rangle' \tag{4.31}$$

Observe that because $|\uparrow\rangle$ and $|\downarrow\rangle$ rotate by $\theta/2$ when the physical basis of the qubit rotates by θ we end up with something very strange when a full $360°$ rotation is performed. This operation maps

$$|\uparrow\rangle \quad\rightarrow\quad -|\uparrow\rangle \quad\text{and}$$
$$|\downarrow\rangle \quad\rightarrow\quad -|\downarrow\rangle$$

This peculiarity is not physically observable, because in the corresponding probability transformation we have a full rotation by $360°$.

Geometric objects that tranform according to equations (4.30) and (4.31) when the physical basis of the qubit is transformed according to equations (4.27) and (4.28) are called *spinors*. We have therefore discovered in this section that the unitary representation of qubits, the qubit Hilbert space, is made of spinors.

Having established this, let us go back to equation (4.29) and work on it some more. In general, a rotation of the canonical basis is described by an orthogonal transformation $\boldsymbol{\Lambda}$ such that

$$
\begin{aligned}
\boldsymbol{e}_{i'} &= \sum_j \Lambda_{i'}{}^j \boldsymbol{e}_j \quad \text{and} \\
\boldsymbol{\omega}^{i'} &= \sum_j \boldsymbol{\omega}^j \Lambda_j{}^{i'}
\end{aligned}
$$

where

$$
\begin{aligned}
\sum_{k'} \Lambda_i{}^{k'} \Lambda_{k'}{}^j &= \delta_i{}^j \quad \text{and} \\
\sum_k \Lambda_{i'}{}^k \Lambda_k{}^{j'} &= \delta_{i'}{}^{j'}
\end{aligned}
$$

We have mentioned this already in section 1.7, page 28. In this more general case equation (4.29) assumes the following form

$$
\boldsymbol{p}' = \frac{1}{2} \begin{pmatrix} 1 + \langle \boldsymbol{\omega}^{z'}, \boldsymbol{r} \rangle \\ 1 - \langle \boldsymbol{\omega}^{z'}, \boldsymbol{r} \rangle \\ 1 + \langle \boldsymbol{\omega}^{x'}, \boldsymbol{r} \rangle \\ 1 + \langle \boldsymbol{\omega}^{y'}, \boldsymbol{r} \rangle \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \sum_i \Lambda_i{}^{z'} \langle \boldsymbol{\omega}^i, \boldsymbol{r} \rangle \\ 1 - \sum_i \Lambda_i{}^{z'} \langle \boldsymbol{\omega}^i, \boldsymbol{r} \rangle \\ 1 + \sum_i \Lambda_i{}^{x'} \langle \boldsymbol{\omega}^i, \boldsymbol{r} \rangle \\ 1 + \sum_i \Lambda_i{}^{y'} \langle \boldsymbol{\omega}^i, \boldsymbol{r} \rangle \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \sum_i \Lambda_i{}^{z'} r^i \\ 1 - \sum_i \Lambda_i{}^{z'} r^i \\ 1 + \sum_i \Lambda_i{}^{x'} r^i \\ 1 + \sum_i \Lambda_i{}^{y'} r^i \end{pmatrix}
$$

Now we have to do something a little confusing. Vectors $\boldsymbol{e}_i$, $i = x, y, z$, and $\boldsymbol{e}_{i'}$, $i' = x', y', z'$, and their dual forms $\boldsymbol{\omega}^i$ and $\boldsymbol{\omega}^{i'}$ operate in the physical 3D space. But in the fiducial space we also have canonical vectors and forms we used to decompose the fiducial vector $\boldsymbol{p}$ into its components. We called them $\boldsymbol{e}_i, i = 0, 1, 2, 3$ and $\boldsymbol{\omega}^i, i = 0, 1, 2, 3$. To avoid a clash of symbols we are going to mark them with a hat:

$$
\hat{\boldsymbol{e}}_i \quad \text{and} \quad \hat{\boldsymbol{\omega}}^i \quad \text{where} \quad i = 0, 1, 2, 3
$$

Moreover, we're going to use primed indexes for them in the expression below, because they refer here to measurements made with respect to the primed directions. Using these we can rewrite our expression for $\boldsymbol{p}'$ as follows:

$$
\begin{aligned}
\boldsymbol{p}' = {} & \frac{1}{2} \Bigg(\left(1 + \sum_{i=x,y,z} \Lambda_i{}^{z'} r^i \right) \hat{\boldsymbol{e}}_{0'} + \left(1 - \sum_{i=x,y,z} \Lambda_i{}^{z'} r^i \right) \hat{\boldsymbol{e}}_{1'} \\
& + \left(1 + \sum_{i=x,y,z} \Lambda_i{}^{x'} r^i \right) \hat{\boldsymbol{e}}_{2'} + \left(1 + \sum_{i=x,y,z} \Lambda_i{}^{y'} r^i \right) \hat{\boldsymbol{e}}_{3'} \Bigg)
\end{aligned}
$$

Now recall equations (2.12), page 63, namely

$$
\begin{aligned}
\hat{\boldsymbol{e}}_{0'} &= \frac{1}{2}\left(\varsigma_{1'} + \varsigma_{z'} - \varsigma_{x'} - \varsigma_{y'}\right) \\
\hat{\boldsymbol{e}}_{1'} &= \frac{1}{2}\left(\varsigma_{1'} - \varsigma_{z'} - \varsigma_{x'} - \varsigma_{y'}\right) \\
\hat{\boldsymbol{e}}_{2'} &= \varsigma_{x'} \\
\hat{\boldsymbol{e}}_{3'} &= \varsigma_{y'}
\end{aligned}
$$

Using these we can replace $\hat{\boldsymbol{e}}_i$ with varsigmas, which yields

$$
\boldsymbol{p}' = \frac{1}{2}\left(\varsigma_{1'} + \sum_{i=x,y,z} \Lambda_i{}^{x'} r^i \, \varsigma_{x'} + \sum_{i=x,y,z} \Lambda_i{}^{y'} r^i \, \varsigma_{y'} + \sum_{i=x,y,z} \Lambda_i{}^{z'} r^i \, \varsigma_{z'}\right)
$$

This tells us that

$$
\begin{aligned}
r^{x'} &= \sum_{i=x,y,z} \Lambda_i{}^{x'} r^i \\
r^{y'} &= \sum_{i=x,y,z} \Lambda_i{}^{y'} r^i \\
r^{z'} &= \sum_{i=x,y,z} \Lambda_i{}^{z'} r^i
\end{aligned}
$$

which is what we knew all along, but we can rewrite this expression differently, grouping the terms around r^x, r^y and r^z instead of $\varsigma_{x'}$, $\varsigma_{y'}$ and $\varsigma_{z'}$ and this time we get:

$$
\boldsymbol{p}' = \frac{1}{2}\left(\varsigma_{1'} + r^x \sum_{i'=x',y',z'} \Lambda_x{}^{i'} \varsigma_{i'} + r^y \sum_{i'=x',y',z'} \Lambda_y{}^{i'} \varsigma_{i'} + r^z \sum_{i'=x',y',z'} \Lambda_z{}^{i'} \varsigma_{i'}\right)
$$

and this tells us that

$$
\varsigma_x = \sum_{i'=x',y',z'} \Lambda_x{}^{i'} \varsigma_{i'} \tag{4.32}
$$

$$
\varsigma_y = \sum_{i'=x',y',z'} \Lambda_y{}^{i'} \varsigma_{i'} \tag{4.33}
$$

$$
\varsigma_z = \sum_{i'=x',y',z'} \Lambda_z{}^{i'} \varsigma_{i'} \tag{4.34}
$$

meaning that the three varsigmas indexed with x, y and z transform like 3D vectors under the rotation of the canonical basis in the qubit's physical space, even though

they have 4 components. The varsigmas are therefore very peculiar objects. They stand with one leg in the qubit's fiducial space and with the other one in the qubit's physical space. They are subject to transformations in both spaces.

To know that they transform like 3D vectors in response to rotations is going to be very useful especially in more complicated situations, where we will have probability matrices that cannot be decomposed into individual probability vectors. Such situations will arise in multi-qubit systems.

4.6   Operators and operands

In the preceding sections of the chapter we managed to recover an image of *pure* states (and pure states *only*) within the formalism of the qubit's Hilbert space that resulted from the *unpacking* of quaternion symbols into 2×2 complex matrices – which, nota bene, all turned out to be Hermitian.

But so far it has only been the density quaternion, or the quaternion equivalent of the probability vector in the fiducial space of a qubit, that we played with. What about the other quaternion, the Hamiltonian?

Recall that

$$\boldsymbol{H} = -\mu \left(B_x \boldsymbol{\sigma}_x + B_y \boldsymbol{\sigma}_y + B_z \boldsymbol{\sigma}_z \right)$$

which translates into

$$\boldsymbol{H} = -\mu \left(\begin{array}{cc} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{array} \right) \tag{4.35}$$

Now, since we can think of $\boldsymbol{H}$ as an *operator*, the obvious question to ask is, what it does to the vectors of Hilbert space, we have identified in the previous sections.

First consider $\boldsymbol{B} = B_z \boldsymbol{e}_z$. For the Hamiltonian constructed from this field

$$\boldsymbol{H}_\uparrow \, |{\uparrow}\rangle = -\mu \left(\begin{array}{cc} B_z & 0 \\ 0 & -B_z \end{array} \right) \left(\begin{array}{c} 1 \\ 0 \end{array} \right) = -\mu B_z \left(\begin{array}{c} 1 \\ 0 \end{array} \right) = -\mu B_z \, |{\uparrow}\rangle$$

Similarly

$$\boldsymbol{H}_\uparrow \, |{\downarrow}\rangle = -\mu \left(\begin{array}{cc} B_z & 0 \\ 0 & -B_z \end{array} \right) \left(\begin{array}{c} 0 \\ 1 \end{array} \right) = \mu B_z \left(\begin{array}{c} 0 \\ 1 \end{array} \right) = \mu B_z \, |{\downarrow}\rangle$$

$|{\uparrow}\rangle$ and $|{\downarrow}\rangle$ are eigenvectors of $\boldsymbol{H}_\uparrow$ and the respective energies are the eigenvalues.

We discover that $|{\uparrow}\rangle$ and $|{\downarrow}\rangle$ are *eigenvectors* of $\boldsymbol{H}_\uparrow$ and the corresponding *eigenvalues* are expectation energies that correspond to these states. Recall that in this case the fiducial formalism would have returned:

$$\langle E_\uparrow \rangle = \langle \boldsymbol{\eta}_\uparrow, \boldsymbol{p}_\uparrow \rangle = \langle -\mu B_z \boldsymbol{\varsigma}^z, \frac{1}{2} \left(\boldsymbol{\varsigma}_1 + \boldsymbol{\varsigma}_z \right) \rangle = -\mu B_z$$

because $\langle \boldsymbol{\varsigma}^z, \boldsymbol{\varsigma}_z \rangle = 2$ and $\langle \boldsymbol{\varsigma}^z, \boldsymbol{\varsigma}_1 \rangle = 0$.

Similarly

$$\langle E_\downarrow \rangle = \langle \boldsymbol{\eta}_\uparrow, \boldsymbol{p}_\downarrow \rangle = \langle -\mu B_z \boldsymbol{\varsigma}^z, \frac{1}{2}(\boldsymbol{\varsigma}_1 - \boldsymbol{\varsigma}_z) \rangle = \mu B_z$$

We have already introduced the notation $|\uparrow\rangle\langle\uparrow|$ and $|\downarrow\rangle\langle\downarrow|$ to describe $\boldsymbol{\rho}_\uparrow$ and $\boldsymbol{\rho}_\downarrow$ as tensor products of unitary vectors and forms, cf. equation (4.10), page 115, and equation (4.11), page 116. In the matrix notation

$$|\uparrow\rangle\langle\uparrow| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad |\downarrow\rangle\langle\downarrow| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

The Hamiltonian $\boldsymbol{H}_\uparrow$ can therefore be described as

$$\boldsymbol{H}_\uparrow = -\mu B_z \, |\uparrow\rangle\langle\uparrow| + \mu B_z \, |\downarrow\rangle\langle\downarrow|$$

Measuring the energy of an individual qubit against this Hamiltonian returns either $-\mu B_z$, and in this case the qubit emerges in the $|\uparrow\rangle$ state from the $\boldsymbol{H}_\uparrow$ measuring apparatus, or μB_z, in which case the qubit emerges in the $|\downarrow\rangle$ from the $\boldsymbol{H}_\downarrow$ measuring apparatus. This we know from the experiment.

From the unitary point of view then, the measurement represented by $\boldsymbol{H}_\uparrow$ per- *Measurements* formed on an individual qubit performs an act of projection. Whatever the state *as projections* of the qubit was originally, the measurement projects the qubit either on the $|\uparrow\rangle$ or on the $|\downarrow\rangle$ state.

The operators $|\uparrow\rangle\langle\uparrow|$ and $|\downarrow\rangle\langle\downarrow|$ may be thought of as projectors: $\boldsymbol{P}_\uparrow$ and $\boldsymbol{P}_\downarrow$. *Hamiltonian* Hence the Hamiltonian itself becomes a linear combination of projections: *eigenstates as*
projection
$$\boldsymbol{H}_\uparrow = -\mu B_z \boldsymbol{P}_\uparrow + \mu B_z \boldsymbol{P}_\downarrow$$
operators

The projectors have the following obvious properties

$$\begin{aligned}
\boldsymbol{P}_\uparrow \boldsymbol{P}_\downarrow &= |\uparrow\rangle\langle\uparrow|\downarrow\rangle\langle\downarrow| = 0 \\
\boldsymbol{P}_\downarrow \boldsymbol{P}_\uparrow &= |\downarrow\rangle\langle\downarrow|\uparrow\rangle\langle\uparrow| = 0 \\
\boldsymbol{P}_\uparrow \boldsymbol{P}_\uparrow &= |\uparrow\rangle\langle\uparrow|\uparrow\rangle\langle\uparrow| = |\uparrow\rangle\langle\uparrow| = \boldsymbol{P}_\uparrow \\
\boldsymbol{P}_\downarrow \boldsymbol{P}_\downarrow &= |\downarrow\rangle\langle\downarrow|\downarrow\rangle\langle\downarrow| = |\downarrow\rangle\langle\downarrow| = \boldsymbol{P}_\downarrow
\end{aligned}$$

Also

$$\boldsymbol{P}_\uparrow + \boldsymbol{P}_\downarrow = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{1}$$

We say that the projectors are orthogonal (because $\boldsymbol{P}_\uparrow \boldsymbol{P}_\downarrow = \boldsymbol{P}_\downarrow \boldsymbol{P}_\uparrow = 0$ – this is what *orthogonal* means) and *idempotent* (because $\boldsymbol{P}_\uparrow \boldsymbol{P}_\uparrow = \boldsymbol{P}_\uparrow$. and $\boldsymbol{P}_\downarrow \boldsymbol{P}_\downarrow = \boldsymbol{P}_\downarrow$ –

this is what *idempotent* means). We can combine these two properties into a single equation, namely

$$\boldsymbol{P}_i \boldsymbol{P}_j = \delta_{ij} \boldsymbol{P}_i$$

where $i \in \{\uparrow, \downarrow\} \ni j$. Also we say that the set of projectors is *complete* (because $\sum_{i \in \{\uparrow,\downarrow\}} \boldsymbol{P}_i = \mathbf{1}$ – this is what *complete* means).

The measurement represented by $\boldsymbol{H}_\uparrow$ defines a complete set of orthogonal projectors. This is true of *any* quantum measurement *within* the unitary formalism.

For an arbitrary superposition $a \ket{\uparrow} + b \ket{\downarrow}$

$$\boldsymbol{H}_\uparrow \left(a \ket{\uparrow} + b \ket{\downarrow} \right) = \left(-\mu B_z \boldsymbol{P}_\uparrow + \mu B_z \boldsymbol{P}_\downarrow \right) \left(a \ket{\uparrow} + b \ket{\downarrow} \right)$$

The first projector $\boldsymbol{P}_\uparrow$ kills $\ket{\downarrow}$, but leaves $\ket{\uparrow}$ intact, because it is a projector, and the second projector $\boldsymbol{P}_\downarrow$ does the opposite. And so the result is

$$-\mu B_z a \ket{\uparrow} + \mu B_z b \ket{\downarrow}$$

$\braket{\Psi | \boldsymbol{H}_\uparrow | \Psi} = \braket{E}$

Now let us zap this result from the left hand side with $\bra{\Psi}$:

$$\begin{aligned}
\left(a^* \bra{\uparrow} + b^* \bra{\uparrow} \right) &\left(a(-\mu B_z) \ket{\uparrow} + b(\mu B_z) \ket{\downarrow} \right) \\
&= aa^*(-\mu B_z) + bb^*(\mu B_z) \\
&= \frac{1}{2} \left(1 + r^z \right) \left(-\mu B_z \right) + \frac{1}{2} \left(1 - r^z \right) \left(\mu B_z \right) \\
&= -\mu B_z r_z \\
&= \braket{E}
\end{aligned}$$

Is this a happy coincidence, or is there more to it? To answer this question we should evaluate $\braket{\Psi | \boldsymbol{H} | \Psi}$ for an arbitrary Hamiltonian $\boldsymbol{H}$ and an arbirary vector $\ket{\Psi}$.

> This transition amplitude, $\braket{\Psi | \boldsymbol{H} | \Psi}$, can be understood as either $\bra{\Psi}$ acting on $\boldsymbol{H} \ket{\Psi}$, or as $\bra{\Psi} \boldsymbol{H}^\dagger$ acting on $\ket{\Psi}$. Both yield the same result, because $\boldsymbol{H}^\dagger = \boldsymbol{H}$.

But rather than plunge into the computation head first, we'll begin by figuring out how individual Pauli matrices affect the basis vectors of the Hilbert space. It is easy to check that:

$$\begin{aligned}
\boldsymbol{\sigma}_x \ket{\uparrow} &= \ket{\downarrow} \\
\boldsymbol{\sigma}_x \ket{\downarrow} &= \ket{\uparrow} \\
\boldsymbol{\sigma}_y \ket{\uparrow} &= i \ket{\downarrow}
\end{aligned}$$

Now we are ready to plunge:

$$\boldsymbol{\sigma}_y \left|\downarrow\right\rangle = -i \left|\uparrow\right\rangle$$
$$\boldsymbol{\sigma}_z \left|\uparrow\right\rangle = \left|\uparrow\right\rangle$$
$$\boldsymbol{\sigma}_z \left|\downarrow\right\rangle = -\left|\downarrow\right\rangle$$

Now we are ready to plunge:

$$-\mu \left(B_x \boldsymbol{\sigma}_x + B_y \boldsymbol{\sigma}_y + B_z \boldsymbol{\sigma}_z\right) \left(a \left|\uparrow\right\rangle + b \left|\downarrow\right\rangle\right)$$
$$= -\mu \left(B_x \left(a \left|\downarrow\right\rangle + b \left|\uparrow\right\rangle\right) + B_y \left(ia \left|\downarrow\right\rangle - ib \left|\uparrow\right\rangle\right) + B_z \left(a \left|\uparrow\right\rangle - b \left|\downarrow\right\rangle\right)\right)$$

Now let us zap it from the left hand side with $a^* \left\langle\uparrow\right| + b^* \left\langle\downarrow\right|$ remembering that $\left\langle\uparrow|\downarrow\right\rangle = \left\langle\downarrow|\uparrow\right\rangle = 0$. This results in:

$$-\mu \left(B_x \left(a^*b + b^*a\right) + B_y i \left(b^*a - a^*b\right) + B_z \left(a^*a - b^*b\right)\right)$$

But recall equations (4.14), (4.15) and (4.13) on page 123:

$$ab^* + ba^* = r^x$$
$$i\left(ab^* - ba^*\right) = r^y$$
$$aa^* - bb^* = r^z$$

Making use of these we get: $\quad\langle \Psi \mid \boldsymbol{H} \mid \Psi \rangle = \langle E \rangle$

$$\langle \Psi \mid \boldsymbol{H} \mid \Psi \rangle = -\mu \left(B_x r^x + B_y r^y + B_z r^z\right) = \langle E \rangle$$

This formula holds for any other Hermitian operator that represents some measurable quantity. Summing up what we have learnt so far about fiducial, quaternion and unitary pictures of quantum systems, we can write: *Expectation values in the unitary formalism*

$$\langle \boldsymbol{\eta}_A, \boldsymbol{p} \rangle = 2\Re \left(\boldsymbol{A\rho}\right) = \mathrm{Tr}\left(\boldsymbol{A\rho}\right) = \langle \Psi \mid \boldsymbol{A} \mid \Psi \rangle = \langle A \rangle$$

Let us still go back to the projection aspect of a measurement on a unitary state. *How the measurement affects the original pure state.* Consider an experiment in which an energy measurement is performed on a state $\left|\Psi\right\rangle = a \left|\uparrow\right\rangle + b \left|\downarrow\right\rangle$. This energy measurement may (and usually does) have a side effect of splitting the incident beam of qubits so that all qubits in one beam emerging from the measuring apparatus are $\left|\uparrow\right\rangle$ and all qubits in the other beam emerging from the measuring apparatus are $\left|\downarrow\right\rangle$. The intensities of both beams would be equal to $I_0 a^* a$ and $I_0 b^* b$ respectively, where I_0 is the intensity of the incident beam.

If we were to mix the two beams back together, we would end up with a *mixture* of qubits in both states in the proportions corresponding to the beams intensities.

Once the measurement has been performed the original pure state, the superposition, is destroyed – in general. But if we were to look at each of the two beams in separation, we'd find that the qubits in them are all in pure states, namely $|\uparrow\rangle$ in one and $|\downarrow\rangle$ in the other, new states induced by the measurement, states that incidentally are the eigenstates of the Hamiltonian.

These states are not $\boldsymbol{P}_\uparrow \mid \Psi\rangle$ or $\boldsymbol{P}_\downarrow \mid \Psi\rangle$, because the projections of $\mid \Psi\rangle$ are not normalized, namely

$$\begin{aligned} \boldsymbol{P}_\uparrow \mid \Psi\rangle &= a \mid \uparrow\rangle \quad \text{and} \\ \boldsymbol{P}_\downarrow \mid \Psi\rangle &= b \mid \downarrow\rangle \end{aligned}$$

How can we express the states that emerge from the measuring apparatus in terms of the original $\mid \Psi\rangle$ and the projectors?

The expectation value of $\boldsymbol{P}_\uparrow$ is

$$\langle\Psi \mid \boldsymbol{P}_\uparrow \mid \Psi\rangle = \langle\Psi \mid a \mid \uparrow\rangle = a^*a\langle\uparrow|\uparrow\rangle = a^*a = |a|^2 = p_\uparrow$$

and similarly for $\boldsymbol{P}_\downarrow$:

$$\langle\Psi \mid \boldsymbol{P}_\downarrow \mid \Psi\rangle = |b|^2 = p_\downarrow$$

In other words, the expectation value of a projection operator $\boldsymbol{P}_i$ is the probability p_i. Therefore

$$\frac{\boldsymbol{P}_\uparrow \mid \Psi\rangle}{\sqrt{\langle\Psi \mid \boldsymbol{P}_\uparrow \mid \Psi\rangle}} = \frac{|a|e^{i\phi_a}}{|a|} \mid \uparrow\rangle = e^{i\phi_a} \mid \uparrow\rangle$$

which is $|\uparrow\rangle$ up to the phase factor $e^{i\phi_a}$. Also

$$\frac{\boldsymbol{P}_\downarrow \mid \Psi\rangle}{\sqrt{\langle\Psi \mid \boldsymbol{P}_\downarrow \mid \Psi\rangle}} = \frac{|b|e^{i\phi_b}}{|b|} \mid \downarrow\rangle = e^{i\phi_b} \mid \downarrow\rangle$$

The effect of the measurement performed by $\boldsymbol{H} = -\mu B_z \boldsymbol{P}_\uparrow + \mu B_z \boldsymbol{P}_\downarrow$ on $\mid \Psi\rangle$ then is to produce two beams in states

$$\frac{\boldsymbol{P}_\uparrow \mid \Psi\rangle}{\sqrt{\langle\Psi \mid \boldsymbol{P}_\uparrow \mid \Psi\rangle}} \quad \text{and} \quad \frac{\boldsymbol{P}_\downarrow \mid \Psi\rangle}{\sqrt{\langle\Psi \mid \boldsymbol{P}_\downarrow \mid \Psi\rangle}} \tag{4.36}$$

with beam intensities equal to $I_0\langle\Psi \mid \boldsymbol{P}_\uparrow \mid \Psi\rangle$ and $I_0\langle\Psi \mid \boldsymbol{P}_\downarrow \mid \Psi\rangle$ respectively.

We can rewrite the above formula in terms of density operators. Let us consider the outcome of the $\boldsymbol{P}_\uparrow$ operation first. Here we find that the state of qubits in the beam produced by this projector is

$$\frac{\boldsymbol{P}_\uparrow \mid \Psi\rangle}{\sqrt{\langle\Psi \mid \boldsymbol{P}_\uparrow \mid \Psi\rangle}} \otimes \frac{\langle\Psi \mid \boldsymbol{P}_\uparrow^\dagger}{\sqrt{\langle\Psi \mid \boldsymbol{P}_\uparrow \mid \Psi\rangle}} = \frac{\boldsymbol{P}_\uparrow \mid \Psi\rangle\langle\Psi \mid \boldsymbol{P}_\uparrow^\dagger}{\langle\Psi \mid \boldsymbol{P}_\uparrow \mid \Psi\rangle} = \frac{\boldsymbol{P}_\uparrow \boldsymbol{\rho} \boldsymbol{P}_\uparrow^\dagger}{\mathrm{Tr}\,(\boldsymbol{P}_\uparrow \boldsymbol{\rho})}$$

Because $\langle \Psi \mid \boldsymbol{P}_\uparrow \mid \Psi \rangle$ and $\mathrm{Tr}\,(\boldsymbol{P}_\uparrow\boldsymbol{\rho})$ are both the same thing, i.e., the expectation value of $\boldsymbol{P}_\uparrow$ on $\boldsymbol{\rho}$.

As the same holds for $\boldsymbol{P}_\downarrow$ we can restate equation (4.36) by saying that the effect of the measurement performed by $\boldsymbol{H} = -\mu B_z\boldsymbol{P}_\uparrow + \mu B_z\boldsymbol{P}_\downarrow$ on $\boldsymbol{\rho}$ is to produce two beams in states

$$\frac{\boldsymbol{P}_\uparrow\boldsymbol{\rho}\boldsymbol{P}_\uparrow^\dagger}{\mathrm{Tr}\,(\boldsymbol{P}_\uparrow\boldsymbol{\rho})} \quad \text{and} \quad \frac{\boldsymbol{P}_\downarrow\boldsymbol{\rho}\boldsymbol{P}_\downarrow^\dagger}{\mathrm{Tr}\,(\boldsymbol{P}_\downarrow\boldsymbol{\rho})} \tag{4.37}$$

Equations (4.37) are more general than equations (4.36), because they can be applied to mixed states as well. The projectors $\boldsymbol{P}_i$, however, are associated with pure states only.

4.7   Properties of the density operator

Suppose we have an arbitrary 2×2 complex matrix $\boldsymbol{\rho}$. We can ask ourselves a question: when is this matrix a plausible density matrix? Being just 2×2 and complex is clearly not enough.

We have answered this question implicitly by developing the whole formalism of quantum mechanics from the probability side rather than from the unitary side, arriving at the expression

$$\boldsymbol{\rho} = \frac{1}{2}\left(\boldsymbol{1} + r^x\boldsymbol{\sigma}_x + r^y\boldsymbol{\sigma}_y + r^z\boldsymbol{\sigma}_z\right) \tag{4.38}$$

where r^x, r^y and r^z are parameters that define the probability vector $\boldsymbol{p}$ in such a way that all its entries are guaranteed to be confined to $[0, 1]$ and the first two entries add up to 1.

In other words, we can say that if $\boldsymbol{\rho}$ can be written in the form (4.38) with r^x, r^y and r^z forming a 3-dimensional real vector of length less than or equal to 1, then $\boldsymbol{\rho}$ is a plausible density matrix.

So, let us rephrase the original question as: "What general conditions does matrix $\boldsymbol{\rho}$ have to satisfy to be rewritable in form (4.38)?

The first condition is obvious. Matrix $\boldsymbol{\rho}$ must be Hermitian. This is because all Pauli matrices are Hermitian. If they weren't Hermitian, they couldn't represent Pauli quaternions and we wouldn't have the mapping from probabilities to 2×2 matrices via quaternions.

Since there are 3 linearly independent Pauli matrices plus the identity matrix, together they constitute a basis in the space of 2×2 complex Hermitian matrices. Every 2×2 Hermitian matrix must be of the form

$$\boldsymbol{\rho} = a\boldsymbol{1} + b\boldsymbol{\sigma}_x + c\boldsymbol{\sigma}_y + d\boldsymbol{\sigma}_z$$

Now, in order to pin a to $1/2$ we can simply request that $\operatorname{Tr}(\boldsymbol{\rho}) = 1$, because all Pauli matrices are traceless and $\operatorname{Tr}(\mathbf{1}) = 2$.

But this still leaves too much freedom to possible values that b, c and d may assume.

Recall that if $\boldsymbol{\rho}$ corresponds to a valid fiducial state $\boldsymbol{p}$ then for every other valid *pure* qubit state $\boldsymbol{p_n}$ defined by some direction $\boldsymbol{n}$, where $\boldsymbol{n} \cdot \boldsymbol{n} = 1$ we have that (cf. equation (4.20), page 127)

$$p_{\boldsymbol{p_n} \leftarrow \boldsymbol{p}} = \langle \tilde{\boldsymbol{p}}_{\boldsymbol{n}}, \boldsymbol{p} \rangle = \langle \tilde{\boldsymbol{p}}_{\boldsymbol{n}} \rangle_{\boldsymbol{p}} \in [0, 1]$$

It is convenient here to switch $\boldsymbol{p_n}$ and $\boldsymbol{p}$ around, so that we consider

$$\langle \tilde{\boldsymbol{p}}, \boldsymbol{p_n} \rangle = \langle \tilde{\boldsymbol{p}} \rangle_{\boldsymbol{p_n}}$$

instead. Since $\boldsymbol{p_n}$ is a pure state we can always find such $\mid \Psi_{\boldsymbol{n}} \rangle$ that the density operator that corresponds to $\boldsymbol{p_n}$ is $\mid \Psi_{\boldsymbol{n}} \rangle \langle \Psi_{\boldsymbol{n}} \mid$. But this is our good old friend, the projector, discussed at some length in section 4.6. Let us call this projector $\boldsymbol{P_n}$ rather than $\boldsymbol{\rho_n}$ then and using this notation we find that

$$\langle \tilde{\boldsymbol{p}} \rangle_{\boldsymbol{p_n}} = \operatorname{Tr}(\boldsymbol{P_n}\boldsymbol{\rho}) = \langle \Psi_{\boldsymbol{n}} \mid \boldsymbol{\rho} \mid \Psi_{\boldsymbol{n}} \rangle$$

Hence we get the following condition

$$\forall_{\boldsymbol{n}} \langle \Psi_{\boldsymbol{n}} \mid \boldsymbol{\rho} \mid \Psi_{\boldsymbol{n}} \rangle \in [0, 1]$$

This condition in a somewhat relaxed form

$$\forall_{\boldsymbol{n}} \langle \Psi_{\boldsymbol{n}} \mid \boldsymbol{\rho} \mid \Psi_{\boldsymbol{n}} \rangle \geq 0 \quad \text{or} \quad \forall_{\boldsymbol{n}} \operatorname{Tr}(\boldsymbol{P_n}\boldsymbol{\rho}) \geq 0 \tag{4.39}$$

is referred to as *positivity* of the density matrix $\boldsymbol{\rho}$.

It is easy to see that this condition is sufficient to enforce $\boldsymbol{r} \cdot \boldsymbol{r} \leq 1$.

Because $\boldsymbol{\rho}$ is Hermitian, it can be diagonalized by "rotating" the basis of the Hilbert space, so that it aligns with the eigenvectors $\mid \eta_i \rangle$ of $\boldsymbol{\rho}$. Strictly speaking "rotations" in the Hilbert space are unitary operations. Now we can choose the eigenvectors $\mid \eta_i \rangle$, $i = 1, 2$ of $\boldsymbol{\rho}$ as some of these $\mid \Psi_{\boldsymbol{n}} \rangle$ vectors and for each of them we still expect that

$$\langle \eta_i \mid \boldsymbol{\rho} \mid \eta_i \rangle = \rho_i \langle \eta_i \mid \eta_i \rangle = \rho_i \in [0, 1]$$

where ρ_i are the eigenvalues of $\boldsymbol{\rho}$. This implies that both eigenvalues of $\boldsymbol{\rho}$ must be confined to $[0, 1]$ too and they must add to 1 as well, otherwise the trace condition

isn't going to be satisfied. The determinant in this case, being a product of both eigenvalues ρ_1 and ρ_2, must be positive.

Because the determinant is an invariant of unitary operations in the Hilbert space, we see that the condition outlined above is captured by

$$\det \boldsymbol{\rho} \geq 0$$

which must hold in any Hilbert space basis.

Recall equation (4.9), page 114. If we allow the r^i coefficients to be anything, it corresponds to our general matrix $\boldsymbol{\rho}$ with $a = 1$, which, as we saw above, derived from the trace condition.

We can now use equation (4.9) to evaluate the determinant:

$$\begin{aligned}
\det \boldsymbol{\rho} &= \frac{1}{4} \left((1 + r^z)(1 - r^z) - (r^x + ir^y)(r^x - ir^y) \right) \\
&= \frac{1}{4} \left(1 - (r^z)^2 - (r^x)^2 - (r^y)^2 \right) \\
&= \frac{1}{4} (1 - \boldsymbol{r} \cdot \boldsymbol{r}) \geq 0
\end{aligned}$$

which implies that

$$\boldsymbol{r} \cdot \boldsymbol{r} \leq 1$$

In summary the following three conditions characterize a plausible density matrix $\boldsymbol{\rho}$

1. Matrix $\boldsymbol{\rho}$ must be Hermitian.

2. The trace of $\boldsymbol{\rho}$ must be 1.

3. Matrix $\boldsymbol{\rho}$ must be positive, i.e.,

$$\forall_{\boldsymbol{n}} \langle \Psi_{\boldsymbol{n}} \mid \boldsymbol{\rho} \mid \Psi_{\boldsymbol{n}} \rangle \geq 0 \quad \text{or} \quad \forall_{\boldsymbol{n}} \operatorname{Tr}(\boldsymbol{P_n \rho}) \geq 0$$

The positivity condition is not an easy condition to use in general, but if we can show that at least one of the eigenvalues of $\boldsymbol{\rho}$ is negative, this disqualifies $\boldsymbol{\rho}$ from being a plausible density matrix right away.

There is one more property that pertains to density operators of pure states only. For such operators we find that

$$\boldsymbol{\rho\rho} = \boldsymbol{\rho} \tag{4.40}$$

because these states are *projectors*.

The proof that $\boldsymbol{\rho}$ of a pure state is idempotent is trivial on the unitary level. Since for a pure state we have that

$$\boldsymbol{\rho} =\mid \Psi \rangle\langle \Psi \mid$$

we can easily see that

$$\boldsymbol{\rho}\boldsymbol{\rho} =\mid \Psi \rangle\langle \Psi \mid \cdot \mid \Psi \rangle\langle \Psi \mid =\mid \Psi \rangle \left(\langle \Psi \mid \Psi \rangle \right) \langle \Psi \mid =\mid \Psi \rangle \left(1 \right) \langle \Psi \mid =\mid \Psi \rangle\langle \Psi \mid = \boldsymbol{\rho}$$

It is instructive to repeat this computation on the quaternion level, because this is going to tell us something about the geometric significance of the density operator's idempotence. Consider $\boldsymbol{\rho} = \left(\mathbf{1} + \sum_i r^i \boldsymbol{\sigma}_i \right)/2$, then

$$\boldsymbol{\rho}\boldsymbol{\rho}$$

$$= \frac{1}{2} \left(\mathbf{1} + \sum_i r^i \boldsymbol{\sigma}_i \right) \cdot \frac{1}{2} \left(\mathbf{1} + \sum_j r^j \boldsymbol{\sigma}_j \right)$$

$$= \frac{1}{4} \left(\mathbf{1} \cdot \mathbf{1} + \mathbf{1} \cdot \sum_j r^j \boldsymbol{\sigma}_j + \left(\sum_i r^i \boldsymbol{\sigma}_i \right) \cdot \mathbf{1} + \sum_i \sum_j r^i r^j \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j \right)$$

$$= \frac{1}{4} \left(\mathbf{1} + 2 \sum_j r^i \boldsymbol{\sigma}_i + \sum_i \sum_j r^i r^j \left(\delta_{ij} \mathbf{1} + i \sum_k \epsilon_{ijk} \boldsymbol{\sigma}_k \right) \right)$$

$$= \frac{1}{4} \left((1 + \boldsymbol{r} \cdot \boldsymbol{r}) \mathbf{1} + 2 \sum_i r^i \boldsymbol{\sigma}_i \right)$$

because $\sum_i \sum_j r^i r^j \epsilon_{ijk} = 0$ on account of $r^i r^j$ being symmetric and ϵ_{ijk} antisymmetric.

Observe that only when $\boldsymbol{r} \cdot \boldsymbol{r} = 1$, i.e., only for pure states, do we end up with

$$\boldsymbol{\rho}\boldsymbol{\rho} = \frac{1}{2} \left(\mathbf{1} + \sum_i r^i \boldsymbol{\sigma}_i \right) = \boldsymbol{\rho}$$

otherwise $\boldsymbol{r}\cdot\boldsymbol{r}$ falls short, it does not reach the 1 and we end up with a deformed quaternion that does not represent any state, because its real component is less than $1/2$.

Conditions 1 through 3 as well as the observation that a density operator of a pure state is idempotent extend to quantum systems of dimensionality higher than those of single qubits, for example to multi-qubit systems and even to infinitely dimensional systems.

4.8   The Schrödinger equation

Like the fiducial vector of a qubit, $\boldsymbol{p}$, and like its corresponding density quaternion $\boldsymbol{\rho}$ a Hilbert space vector $\mid \Psi \rangle$, that describes a pure qubit state, evolves too in the

presence of a "magnetic field" $\boldsymbol{B}$. This evolution can be easily derived by taking apart the von Neumann equation (3.3) discussed in section 3.6 on page 100:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\rho} = -\frac{i}{\hbar}\,[\boldsymbol{H},\boldsymbol{\rho}]$$

The trick is to substitute

$$\boldsymbol{\rho} = \mid\Psi\rangle\langle\Psi\mid$$

Unpacking the von Neumann equation

We don't even have to split $\mid\Psi\rangle$ into a superposition. Consider first the left hand side of the resulting equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\rho} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mid\Psi\rangle\langle\Psi\mid\right) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\mid\Psi\rangle\right)\langle\Psi\mid + \mid\Psi\rangle\left(\frac{\mathrm{d}}{\mathrm{d}t}\langle\Psi\mid\right)$$

On the right hand side we have

$$-\frac{i}{\hbar}\left(\boldsymbol{H}\mid\Psi\rangle\langle\Psi\mid - \mid\Psi\rangle\langle\Psi\mid\boldsymbol{H}\right)$$

Combining the two we discover a sum of two equations:

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\mid\Psi\rangle\right)\langle\Psi\mid \;=\; -\frac{i}{\hbar}\left(\boldsymbol{H}\mid\Psi\rangle\right)\langle\Psi\mid \quad \text{and}$$

$$\mid\Psi\rangle\left(\frac{\mathrm{d}}{\mathrm{d}t}\langle\Psi\mid\right) \;=\; \mid\Psi\rangle\frac{i}{\hbar}\left(\langle\Psi\mid\boldsymbol{H}\right)$$

These two equations are duals of each other and they reduce to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mid\Psi\rangle = -\frac{i}{\hbar}\boldsymbol{H}\mid\Psi\rangle \tag{4.41}$$

This is the celebrated Schrödinger equation for a qubit. Like the von Neumann equation and its fiducial space equivalent this equation preserves the purity of the state. We call this evolution *unitary* because it does not affect the length of the state vector $\mid\Psi\rangle$, which remains $\langle\Psi\mid\Psi\rangle = 1$, or... *unity*. The name "unitary" is also used to describe the whole formalism that results from unpacking quaternions into Pauli matrices and that represents quantum states by vectors in the Hilbert space on which the matrices operate, rather than by density quaternions (or operators). This formalism is the subject of this chapter.

Schrödinger equation

Observe that $\boldsymbol{H}$ is the most general 2×2 *Hermitian* matrix. For such a matrix the diagonal elements must be real, because in this case $H_{ij} = H_{ji}^*$ implies $H_{ii} = H_{ii}^*$, and off diagonal elements must be complex conjugates of

Matrix $\boldsymbol{H}$ is universal – it describes all 2-dimensional quantum systems

their mirror images across the diagonal. This leaves us with 4 independent parameters and we find them all here in equation (4.35). Well, we actually find only 3, but remember that $\boldsymbol{H}$ represents energy, and energy is only defined up to an additive constant. We can therefore always choose the constant so that $H_{11} = -H_{22}$.

A Hamiltonian is responsible for a unitary evolution of a quantum system. Because $\boldsymbol{H}$ is the most general 2×2 Hamiltonian possible this means that *any* other quantum system characterized by two basis states only must be described by a matrix that looks the same. The interpretation of vector $\boldsymbol{B}$, of course, differs from a system to a system, as does the intepretation of $\boldsymbol{r}$, but the equations and their solutions are identical.

This is why whether we talk about a quantronium or about a neutron spin, two systems that couldn't be more different at first glance, we end up with exactly the same mathematics, the same dynamics and the same properties. Whatever can be said about a neutron beam, translates immediately into pronouncements that can be made about a statistical ensemble of quantroniums, or two level molecules, or two level quantum dots, or any other two-level quantum system.

4.8.1   General solution of the Schrödinger equation

Equation (4.41) can be solved quite easily for a general case of $\boldsymbol{H}$. We are going to solve it first for $\boldsymbol{H} = \text{constant}$.

Consider a simple discrete approximation of the Schrödinger equation:

$$\frac{\mathrm{d}\boldsymbol{H}}{\mathrm{d}t} \approx \frac{\mid \Psi(t+\Delta t)\rangle - \mid \Psi(t)\rangle}{\Delta t} = -\frac{i}{\hbar}\boldsymbol{H} \mid \Psi(t)\rangle$$

Advancing the Schrödinger equation by a single Euler time step

We can extract $\mid \Psi(t+\Delta t)\rangle$ from it and this leads to the familiar Euler time step:

$$\mid \Psi(t+\Delta t)\rangle = \mid \Psi(t)\rangle - \frac{i}{\hbar}\boldsymbol{H} \mid \Psi(t)\rangle\Delta t$$

This equation tells us something quite insightful about the Schrödinger equation to begin with. The Schrödinger equation basically represents the simplest evolution possible. It says that evolved over a short time span Δt vector $\mid \Psi(t+\Delta t)\rangle$ is going to differ from the original vector $\mid \Psi(t)\rangle$ by a small linear correction $-\frac{i}{\hbar}\boldsymbol{H} \mid \Psi(t)\rangle\Delta t$. There are no fancy second derivatives here as we have in the Newton's equations, no third derivatives as we have in the Lorentz-Abraham equations, no complicated curvature terms and connection symbols as we have in the Einstein's equations of General Relativity. It is amazingly simple.

Let us use this insight to figure out how a quantum system is going to evolve over a longer time stretch.

Let us take $t = 0$ as our starting point. After a sufficiently short Δt an initial state $\mid \Psi(0)\rangle$ will evolve into

$$\mid \Psi(\Delta t)\rangle = \left(\mathbf{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)\mid \Psi(0)\rangle$$

Having made this one time step, we can make another one, also of length Δt:

$$
\begin{aligned}
\mid \Psi(2\Delta t)\rangle &= \left(\mathbf{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)\mid \Psi(\Delta t)\rangle \\
&= \left(\mathbf{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)\left(\mathbf{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)\mid \Psi(0)\rangle
\end{aligned}
$$

It is now clear that for every additional time step of length Δt, we're going to act on the initial state $\mid \Psi(0)\rangle$ with a yet another instance of $\left(\mathbf{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)$. In summary:

$$\mid \Psi(n\Delta t)\rangle = \left(\mathbf{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)^n\mid \Psi(0)\rangle$$

But $n\Delta t = t$, so

$$\mid \Psi(t)\rangle = \left(\mathbf{1} + \frac{1}{i\hbar}\boldsymbol{H}\frac{t}{n}\right)^n\mid \Psi(0)\rangle$$

Advancing the Schrödinger equation by multiple Euler time steps

This expression is approximate, because it results from taking n finite, though small, time steps of length $\Delta t = t/n$. Clearly, we can only get more accurate by making $\Delta t = t/n$ smaller, which is the same as taking n larger, converging on the exact solution for $n \to \infty$:

$$\mid \Psi(t)\rangle = \lim_{n\to\infty}\left(\mathbf{1} + \frac{1}{i\hbar}\boldsymbol{H}\frac{t}{n}\right)^n\mid \Psi(0)\rangle$$

This limit can be evaluated as follows. Recall the familiar Newton formula for $(a + b)^n$:

$$(a + b)^n = \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k$$

We can use this formula here, because $\mathbf{1}\boldsymbol{H} = \boldsymbol{H}\mathbf{1}$ (the derivation of the Newton formula makes use of $ab = ba$), and this yields

$$\left(\mathbf{1} + \frac{\boldsymbol{H}t}{i\hbar n}\right)^n = \sum_{k=0}^{n}\binom{n}{k}\mathbf{1}^{n-k}\left(\frac{\boldsymbol{H}t}{i\hbar n}\right)^k$$

Of course $\mathbf{1}$ applied any number of times is still $\mathbf{1}$, so we can drop it. We can also unpack the $\binom{n}{k}$ symbol and this is what we end up with:

$$\left(\mathbf{1} + \frac{\boldsymbol{H}t}{i\hbar n}\right)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \left(\frac{\boldsymbol{H}t}{i\hbar n}\right)^k$$

Let us have a closer look at $n!/(k!(n-k)!)$. This can be rewritten as:

$$\frac{(n-k+1)(n-k+2)\ldots(n-k+k)}{k!}$$

We have clearly k terms in the numerator and they are all of the form $(n - \text{something})$ with the exception of the last one, which is just n. If we were to evaluate this we'd get

$$n^k + n^{k-1} \times \text{something} + n^{k-2} \times \text{something else} + \ldots$$

But observe that there is also n^k in the denominator of $(\boldsymbol{H}t/(i\hbar n))^k$. For $n \to \infty$ $n^k/n^k = 1$, but all the other terms like $n^{k-1} \times \text{something}/n^k$ become zero. Consequently we end up with:

$$\lim_{n\to\infty} \left(\mathbf{1} + \frac{\boldsymbol{H}t}{i\hbar n}\right)^n = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{\boldsymbol{H}t}{i\hbar}\right)^k \tag{4.42}$$

Ah, but this looks so much like e^x, recall that

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{4.43}$$

Sure, we have just a plain number x in (4.43), but an operator $\boldsymbol{H}t/(i\hbar)$ in (4.42). We know that $e^n = e \times e \times \ldots \times e$ n-times. But what does $e^{\boldsymbol{H}}$ mean? Well, it means:

$$e^{-i\boldsymbol{H}t/\hbar} \doteq \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{\boldsymbol{H}t}{i\hbar}\right)^k$$

General solution to the Schrödinger equation with a constant Hamiltonian

This is how we *define* it and having done so we can write the general solution to the Schrödinger equation for $\boldsymbol{H} = \text{constant}$ as:

$$|\Psi(t)\rangle = e^{-i\boldsymbol{H}t/\hbar} |\Psi(0)\rangle \tag{4.44}$$

Properties of matrix exponential

The matrix exponential $e^{\boldsymbol{A}}$ does not always have the same properties as an ordinary number exponential e^x. The reason for this is that matrices

do not commute in general, whereas numbers (with the notable exception of quaternions) do. In particular we cannot always write

$$e^{\boldsymbol{A}} e^{\boldsymbol{B}} = e^{\boldsymbol{A}+\boldsymbol{B}}$$

This equation applies only when $\boldsymbol{A}$ and $\boldsymbol{B}$ commute, i.e., when $\boldsymbol{AB} = \boldsymbol{BA}$. Of course, since $\boldsymbol{A}$ commutes with itself, we can always write

$$e^{a\boldsymbol{A}} e^{b\boldsymbol{A}} = e^{(a+b)\boldsymbol{A}}$$

But other properties of the exponential, such as $e^0 = 1$ and $e^{-x} = 1/e^x$ still hold with appropriate matrix substitutions for the inverse and for the 1:

$$\begin{aligned} e^{\boldsymbol{0}} &= \boldsymbol{1} \\ \left(e^{\boldsymbol{A}}\right)^{-1} &= e^{-\boldsymbol{A}} \end{aligned}$$

The other very useful property is that when $\boldsymbol{A}$ and $\boldsymbol{B}$ are *similar*, i.e., such that

$$\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$$

where $\boldsymbol{M}$ is an arbitrary invertible matrix (sized so that the equation above makes sense) then

$$e^{\boldsymbol{B}} = \boldsymbol{M}^{-1} e^{\boldsymbol{A}} \boldsymbol{M}$$

This is very easy to see when you consider that

$$\boldsymbol{B}\boldsymbol{B}\dots\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}\dots\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{A}\dots\boldsymbol{A}\boldsymbol{M}$$

The reason why this is such a useful property is that it lets us evaluate $\exp\boldsymbol{B}$ easily if $\boldsymbol{B}$ is diagonalizable, e.g., to $\boldsymbol{A}$, because then $\exp\boldsymbol{A}$ is a diagonal matrix filled with exponentials of the diagonal terms of $\boldsymbol{A}$.

But what if $\boldsymbol{H}$ is not constant, what if $\boldsymbol{H} = \boldsymbol{H}(t)$?

The problem is still tractable. Much depends on how $\boldsymbol{H}$ varies with t. Suppose $\boldsymbol{H}$ is constant and equal to, say, $\boldsymbol{H}_1$ for $t \in [0, t_1]$, then $\boldsymbol{H}$ may change rapidly to $\boldsymbol{H}_2$ for $t \in\ ]t_1, t_2]$, and so on. The solution in this case will be

General solution to the Schrödinger equation with time dependent Hamiltonian

$$\begin{aligned} |\,\Psi(t)\rangle &= e^{-i\boldsymbol{H}_1 t/\hbar}\,|\,\Psi(0)\rangle \quad \text{for} \quad t \in [0, t_1] \\ |\,\Psi(t)\rangle &= e^{-i\boldsymbol{H}_2(t-t_1)/\hbar} e^{-i\boldsymbol{H}_1 t_1/\hbar}\,|\,\Psi(0)\rangle \quad \text{for} \quad t \in\ ]t_1, t_2] \\ |\,\Psi(t)\rangle &= e^{-i\boldsymbol{H}_3(t-t_2)/\hbar} e^{-i\boldsymbol{H}_2(t_2-t_1)/\hbar} e^{-i\boldsymbol{H}_1 t_1/\hbar}\,|\,\Psi(0)\rangle \quad \text{for} \quad t \in\ ]t_2, t_3] \\ &\dots \end{aligned}$$

and so on.

Consider the Schrödinger equation with a time dependent Hamiltonian

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t} \mid \Psi(t)\rangle = \boldsymbol{H}(t) \mid \Psi(t)\rangle$$

Let us chop time into small segments Δt and let us assume that the Hamiltonian changes sufficiently slowly so that we can consider it constant within each segment. Following the reasoning presented above we can write the solution of this equation in the following form:

$$
\begin{aligned}
\mid \Psi(t)\rangle \quad = \quad & e^{-i\boldsymbol{H}(t)\Delta t/\hbar}e^{-i\boldsymbol{H}(t-\Delta t)\Delta t/\hbar}e^{-i\boldsymbol{H}(t-2\Delta t)\Delta t/\hbar} \\
& \dots e^{-i\boldsymbol{H}(\Delta t)\Delta t/\hbar}e^{-i\boldsymbol{H}(0)\Delta t/\hbar} \mid \Psi(0)\rangle \qquad (4.45)
\end{aligned}
$$

Of course, $e^{\boldsymbol{H}_1}e^{\boldsymbol{H}_2} \neq e^{\boldsymbol{H}_1+\boldsymbol{H}_2}$, if $\boldsymbol{H}_1$ and $\boldsymbol{H}_2$ do not commute. But suppose that the evolution of $\boldsymbol{H}(t)$ is such that $[\boldsymbol{H}(t_1), \boldsymbol{H}(t_2)] = \boldsymbol{0}$ for each t_1 and t_2. In this happy case we are allowed to gather all the exponents into a sum:

$$\mid \Psi(t)\rangle = e^{-i(\boldsymbol{H}(t)+\boldsymbol{H}(t-\Delta t)+\boldsymbol{H}(t-2\Delta t)+\dots+\boldsymbol{H}(\Delta t)+\boldsymbol{H}(0))\Delta t/\hbar} \mid \Psi(0)\rangle$$

The shorter the Δt the more accurate the expression, so in the limit of $\Delta t \to 0$ we get

$$\mid \Psi(t)\rangle = e^{-i\left(\int_0^t \boldsymbol{H}(t)\,\mathrm{d}t\right)/\hbar} \mid \Psi(0)\rangle \qquad (4.46)$$

Unitarity of the Schrödinger evolution

Expression such as $\exp\left(-i\boldsymbol{H}\Delta t/\hbar\right)$ represents a finite transformation of a quantum system enacted by $\boldsymbol{H}$ that was applied to the system for the duration of Δt. It characteristic feature is that it does not affect the length of vector $\Psi(t)$:

$$\langle\Psi(\Delta t) \mid \Psi(\Delta t)\rangle = \langle\Psi(0) \mid e^{i\boldsymbol{H}^\dagger\Delta t/\hbar}e^{-i\boldsymbol{H}\Delta t/\hbar} \mid \Psi(0)\rangle$$

Recall that $\boldsymbol{H}$ is Hermitian, so that $\boldsymbol{H}^\dagger = \boldsymbol{H}$, hence

$$e^{i\boldsymbol{H}^\dagger\Delta t/\hbar}e^{-i\boldsymbol{H}\Delta t/\hbar} = e^{i\boldsymbol{H}\Delta t/\hbar}e^{-i\boldsymbol{H}\Delta t/\hbar} = e^{i(\boldsymbol{H}-\boldsymbol{H})\Delta t/\hbar} = e^{\boldsymbol{0}} = \boldsymbol{1}$$

We could gather the exponents into a sum, because $\boldsymbol{H}$ commutes with itself. And so in the end we get

$$\langle\Psi(\Delta t) \mid \Psi(\Delta t)\rangle = \langle\Psi(0) \mid \Psi(0)\rangle \qquad (4.47)$$

A most general such operation, given by equation (4.45), is a superposition of unitary operations and thus a unitary operation itself. Let us call it $\boldsymbol{U}(t)$, so that

$$\mid \Psi(t)\rangle = \boldsymbol{U}(t) \mid \Psi(0)\rangle$$

Its dual equivalent, $\boldsymbol{U}^\dagger(t)$, given by

$$
\begin{aligned}
\boldsymbol{U}^\dagger(t) \quad = \quad & e^{i\boldsymbol{H}(0)\Delta t/\hbar}e^{i\boldsymbol{H}(\Delta t)\Delta t/\hbar} \\
& \dots e^{i\boldsymbol{H}(t-2\Delta t)\Delta t/\hbar}e^{i\boldsymbol{H}(t-\Delta t)\Delta t/\hbar}e^{i\boldsymbol{H}(t)\Delta t/\hbar}
\end{aligned}
$$

evolves a form that is dual to vector $\mid \Psi \rangle$:

$$\langle \Psi(t) \mid = \langle \Psi(0) \mid \boldsymbol{U}^\dagger(t)$$

When put together, they annihilate each other from the middle onwards:

$$
\begin{aligned}
\langle \Psi(t) \mid \Psi(t) \rangle &= \langle \Psi(0) \mid \boldsymbol{U}^\dagger(t)\boldsymbol{U}(t) \mid \Psi(0) \rangle \\
&= \langle \Psi(0) \mid e^{i\boldsymbol{H}(0)\Delta t/\hbar}e^{i\boldsymbol{H}(\Delta t)\Delta t/\hbar} \\
&\quad \dots \left(e^{i\boldsymbol{H}(t-2\Delta t)\Delta t/\hbar} \left(e^{i\boldsymbol{H}(t-\Delta t)\Delta t/\hbar} \right.\right. \\
&\quad \left(e^{i\boldsymbol{H}(t)\Delta t/\hbar}e^{-i\boldsymbol{H}(t)\Delta t/\hbar}\right) \\
&\quad \left.\left. e^{-i\boldsymbol{H}(t-\Delta t)\Delta t/\hbar} \right)e^{-i\boldsymbol{H}(t-2\Delta t)\Delta t/\hbar} \right) \\
&\quad \dots e^{-i\boldsymbol{H}(\Delta t)\Delta t/\hbar}e^{-i\boldsymbol{H}(0)\Delta t/\hbar} \mid \Psi(0) \rangle \\
&= \langle \Psi(0) \mid \Psi(0) \rangle
\end{aligned}
$$

A qubit evolution operator $\boldsymbol{U}(t)$ can be represented by a 2×2 complex matrix. These matrices have the following property:

$$\boldsymbol{U}(t)\boldsymbol{U}^\dagger(t) = \boldsymbol{1}, \tag{4.48}$$

which we have just demonstrated. Matrices that satisfy this property are called *unitary*, and the corresponding operators are called *unitary operators*. Hence the name of the formalism.

Another way to look at the unitary operators is to observe that

$$\boldsymbol{U}^\dagger = \boldsymbol{U}^{-1},$$

that is, the Hermitian conjugate of $\boldsymbol{U}$ is its inverse.

Unitary operators are closely related to orthogonal operators, i.e., rotations and reflections. Indeed, if $\boldsymbol{U}$ is real, then $\boldsymbol{U}^\dagger = \boldsymbol{U}^T$ and equation (4.48) becomes

$$\boldsymbol{U}(t)\boldsymbol{U}^T(t) = \boldsymbol{1},$$

which defines orthogonal operators.

A combination of $\boldsymbol{U}(t)$ and $\boldsymbol{U}^\dagger(t)$ is needed to evolve a density operator made of $\mid \Psi \rangle$ and $\langle \Psi \mid$:

$$
\begin{aligned}
\boldsymbol{\rho}(t) &= \mid \Psi(t) \rangle\langle \Psi(t) \mid \\
&= \boldsymbol{U}(t) \mid \Psi(0) \rangle\langle \Psi(0) \mid \boldsymbol{U}^\dagger(t) \\
&= \boldsymbol{U}(t)\boldsymbol{\rho}(0)\boldsymbol{U}^\dagger(t) \tag{4.49}
\end{aligned}
$$

It is easy to get from here back to the von Neumann equation. Consider a short-time-increment version of $\boldsymbol{U}$, that is,

$$\boldsymbol{U}(\Delta t) \;=\; \boldsymbol{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t, \quad \text{and}$$

$$\boldsymbol{U}^{\dagger}(\Delta t) \;=\; \boldsymbol{1} - \frac{1}{i\hbar}\boldsymbol{H}\Delta t,$$

where, again, we have made use of $\boldsymbol{H} = \boldsymbol{H}^{\dagger}$. Applying these to $\boldsymbol{\rho}(0)$ yields

$$\boldsymbol{\rho}(\Delta t) = \boldsymbol{U}(\Delta t)\boldsymbol{\rho}(0)\boldsymbol{U}^{\dagger}(\Delta t)$$

$$= \left(\boldsymbol{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)\boldsymbol{\rho}(0)\left(\boldsymbol{1} - \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)$$

$$= \left(\boldsymbol{\rho}(0) + \frac{\Delta t}{i\hbar}\boldsymbol{H}\boldsymbol{\rho}(0)\right)\left(\boldsymbol{1} - \frac{1}{i\hbar}\boldsymbol{H}\Delta t\right)$$

$$= \boldsymbol{\rho}(0) + \frac{\Delta t}{i\hbar}\boldsymbol{H}\boldsymbol{\rho}(0) - \frac{\Delta t}{i\hbar}\boldsymbol{\rho}(0)\boldsymbol{H} + \mathcal{O}(\Delta t)^2$$

$$\approx \boldsymbol{\rho}(0) + \frac{1}{i\hbar}\left[\boldsymbol{H}, \boldsymbol{\rho}(0)\right]$$

And this implies

$$\frac{\boldsymbol{\rho}(\Delta t) - \boldsymbol{\rho}(0)}{\Delta t} = \frac{1}{i\hbar}\left[\boldsymbol{H}, \boldsymbol{\rho}(0)\right],$$

which is a finite difference approximation of the von Neumann equation.

The 2×2 unitary operators $\boldsymbol{U}$ and $\boldsymbol{U}^{\dagger}$ are not Hermitian. Nevertheless, they can be represented in terms of Pauli matrices, and mapped onto quaternions, assuming that some of the coefficients are complex. For example, the small Δt form of $\boldsymbol{U}$ for a magnetized qubit is given by

$$\boldsymbol{U} \;=\; \boldsymbol{1} + \frac{1}{i\hbar}\boldsymbol{H}\Delta t$$

$$= \boldsymbol{1} - \frac{\mu\Delta t}{i\hbar}\left(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z\right)$$

Together with $\boldsymbol{U}^{\dagger}$ they produce a small rotation of the polarization vector $\boldsymbol{r}$ by $\omega_L\Delta t$ about the direction of vector $\boldsymbol{B}$, where $\omega_L = \frac{2\mu B}{\hbar}$ and $B = \sqrt{B_x^2 + B_y^2 + B_z^2}$. Each by itself produces a corresponding small "rotation" of the unitary equivalent of $\boldsymbol{r}$, either $\mid\Psi\rangle$ or $\langle\Psi\mid$, in the spinor space. Their compositions, $(\boldsymbol{U}(\Delta t))^n$, are equivalent to multiple small rotations of $\boldsymbol{r}$, and add up to large finite rotations of $\boldsymbol{r}$.

All that the unitary machinery of quantum mechanics can do to a single qubit is to rotate its polarization vector $\boldsymbol{r}$.

4.8.2 Larmor precession revisited

Suppose $| \Psi_n(0) \rangle$ for $n = 1, 2$ are eigenstates of $\boldsymbol{H}$. In this case

$$\boldsymbol{H} \mid \Psi_n \rangle = E_n \mid \Psi_n \rangle$$

Applying the Schrödinger equation to the eigenstates results in:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \mid \Psi_n(t) \rangle = E_n \mid \Psi_n(t) \rangle$$

which has the simple solution

$$\mid \Psi_n(t) \rangle = e^{-iE_n t/\hbar} \mid \Psi_n(0) \rangle$$

Hamiltonian eigenstates "vibrate" with eigenfrequency

This is a special case of the general solution given by equation (4.44) on page 148.

Each of the eigenstates appears to "vibrate" with its own eigenfrequency

$$\omega_n = E_n/\hbar$$

The corresponding forms vibrate with the same frequencies but in the opposite direction:

$$\langle \Psi_n(t) \mid = \langle \Psi_n(0) \mid e^{iE_n t/\hbar}$$

In all expressions of the kind $\langle \Psi \mid \boldsymbol{H} \mid \Psi \rangle$ the two vibrations, that of the vector and that of the form, cancel. Similarly, if you look at the density operator $\boldsymbol{\rho} = \mid \Psi \rangle \langle \Psi \mid$ the vibration terms cancel again, so the "vibrations" of the eigenvectors are physically unobservable. The eigenstates just stay put and don't change. In order to force a change, e.g., to make the qubit flip, we have to use a different Hamiltonian $\boldsymbol{H}_1$, such that the original eigenstates of $\boldsymbol{H}$ are no longer the eigenstates of $\boldsymbol{H}_1$.

Undetectability of eigenfrequencies

But now consider a superposition of $|\uparrow\rangle$ and $|\downarrow\rangle$, for example

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle)$$

and assume that $\boldsymbol{H} = -\mu B_z \boldsymbol{\sigma}_z$. Both $|\uparrow\rangle$ and $|\downarrow\rangle$ are the eigenstates of this Hamiltonian, as we saw in section 4.6, page 136.

Superpositions of Hamiltonian eigenstates precess with Larmor frequency, which corresponds to the difference between eigenenergies.

Each of the two eigenstates will evolve in its own way, independent of the other one, because the Schrödinger equation is linear. If $| \Psi(0) \rangle = |\rightarrow\rangle$ then at some later time t

$$\mid \Psi(t) \rangle = \frac{1}{\sqrt{2}} \left(e^{i\mu B_z t/\hbar} |\uparrow\rangle + e^{-i\mu B_z t/\hbar} |\downarrow\rangle \right)$$

Let us invoke again equations (4.13), (4.14) and (4.15) from page 122, namely:

$$
\begin{aligned}
aa^* - bb^* &= r^z \\
ab^* + ba^* &= r^x \\
i\left(ab^* - ba^*\right) &= r^y
\end{aligned}
$$

Here $a = \frac{1}{\sqrt{2}}e^{i\mu B_z t/\hbar}$ and $b = \frac{1}{\sqrt{2}}e^{-i\mu B_z t/\hbar}$. The first thing we notice is that $r^z = 0$, so $\boldsymbol{r}$ remains in the plane perpendicular to $\boldsymbol{e}_z$. But

$$
\begin{aligned}
r^x &= \frac{1}{2}\left(e^{i2\mu B_z t/\hbar} + e^{-i2\mu B_z t/\hbar}\right) = \cos\frac{2\mu B_z t}{\hbar} \quad \text{and} \\
r^y &= \frac{i}{2}\left(e^{i2\mu B_z t/\hbar} - e^{-i2\mu B_z t/\hbar}\right) = -\sin\frac{2\mu B_z t}{\hbar}
\end{aligned}
$$

We see that state $|\rightarrow\rangle$ precesses about the z axis with Larmor frequency

$$
\omega_L = \frac{2\mu B_z}{\hbar}
$$

In summary, even though the so called *phase factors*, with which the eigenstates "vibrate", namely $\exp\left(-iE_n t/\hbar\right)$, are invisible in isolation, they become detectable in superpositions, where they manifest as Larmor precession.

Unitary formalism hides three-dimensional character of qubits' kinematics and dynamics

Because our dynamic equation that described the evolution of qubit probabilities, and then its quaternion equivalent, the von Neumann equation that described the evolution of the density quaternion were restricted to fully polarized states and did not describe depolarization, they were, in fact, fully equivalent to the Schrödinger equation. The unitary formalism reproduces all that we covered in our discussion of qubit dynamics. The difference is that the unitary formalism *hides* vector $\boldsymbol{r}$ inside the two *complex* coefficients a and b that multiply the two basis vectors of the Hilbert space. People who look at superpositions such as $a\,|\uparrow\rangle + b\,|\downarrow\rangle$ often think of a and b as two *real* numbers and apply intuitions that pertain to *real* vector spaces. But even with the normalization condition $aa^* + bb^* = 1$ imposed, there are, in fact, three real numbers hidden inside a and b, not just two, and together they encode the three-dimensionality of qubit kinematics and dynamics.

4.9   Single qubit gates

The highly suggestive qubit notation employed by the unitary formalism lets us identify qubit states $|\uparrow\rangle$ and $|\downarrow\rangle$ with 0 and 1 of Boolean logic. But, of course, $|\uparrow\rangle$ and $|\downarrow\rangle$ are not 0 and 1. Generally, a qubit state $|\,\Psi\rangle$ corresponds to a 3-dimensional

vector of length 1 (or of length that is no greater than 1 if we allow for mixtures) that can point in any direction. This, as we have emphasized above, may not be clear within the confines of the unitary formalism, because r is hidden inside a, a^*, b and b^*, but it is there.

The Rabi oscillations discussed in chapter 2, section 2.10 on page 75 and then illustrated with the example of a quantum circuit, the quantronium, discussed in section 2.11 (page 83) provided us with an example of controllable driven evolution *Driving qubit* of a qubit that did not result in depolarization – at least on paper, we had actually *evolution* seen depolarization in the quantronium example. The tip of the qubit's vector r drew a continuous line on the Bloch sphere in the course of the evolution. The Rabi oscillations were slow compared to the Larmor precession. This was what made them controllable and precise. For this reason the Rabi oscillations are a preferred method for executing various computational operations on qubits. But it is not impossible to use the controlled Larmor precession for this purpose, at least in principle. By lowering the value of the magnetic field $\boldsymbol{B}_\parallel$ we can slow down the pace of the Larmor precession to the point where it can be controllable.

The simplest logical operation is the NOT gate. We have analyzed it already in *The* NOT *gate* section 2.10. But let us rehash the general idea here. Assume the qubit is in the $|\uparrow\rangle \equiv |\,0\rangle$ state originally and sits in the strong background field $\boldsymbol{B}_\parallel$ – storing its computational value. In order to flip the qubit from $|\,0\rangle$ to $|\,1\rangle$ we can buzz it with $\boldsymbol{B}_\perp = B_\perp \left(\sin \omega_L t \boldsymbol{e}_x - \cos \omega_L t \boldsymbol{e}_y\right)$, where $\omega_L = 2\mu B_\parallel/\hbar$, for $\pi\hbar/(2\mu B_\perp)$ seconds exactly. If the initial state of the qubit was $|\downarrow\rangle \equiv |\,1\rangle$ the same operation would flip it to $|\,0\rangle$. This is what makes this operation a proper NOT gate: it has to do the right thing for both $|\,0\rangle$ and $|\,1\rangle$ at the same time.

We can draw the following quantum circuit representations for the NOT gate. *Diagrammatic representation of the* NOT *gate*
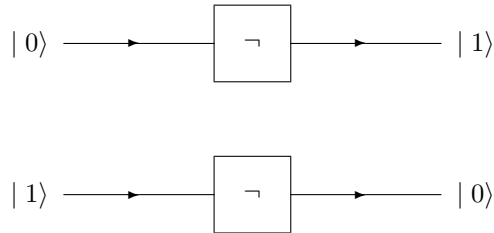

Figure 4.2: Diagrammatic representation of the NOT gate.

The lines with arrows symbolize, say, a polarized neutron beam. The box la-

beled with the logical NOT symbol, $\neg$, stands for, say, a chamber filled with the combination of $\boldsymbol{B}_\perp$ and $\boldsymbol{B}_\parallel$ needed to perform the operation. The dimensions of the chamber can be set so that, given the beam's velocity v, the neutrons of the beam would spend exactly the right amount of time in the chamber. We have used the word *say* in order to emphasize that what the symbols drawn in figure 4.2 represent, depends on the particular implementation of the qubit.

Arithmetic definition of NOT

The quantum NOT operation can be defined arithmetically too

$$\neg \,|\,0\rangle \;=\; |\,1\rangle \tag{4.50}$$

$$\neg \,|\,1\rangle \;=\; |\,0\rangle \tag{4.51}$$

In the world of classical physics this wouldn't be enough to specify the operation uniquely. Every rotation by $180°$ about an axis perpendicular to $\boldsymbol{e}_z$ could be used to implement the operation and there is an infinite number of such axes bisecting the great circle of the Bloch sphere. But in the peculiar world of quantum physics $|\,0\rangle$ and $|\,1\rangle$ are the physicsl basis states and basis states in the Hilbert space of a qubit as well. This means that equations (4.50) and (4.51) should also apply to superpositions of $|\,0\rangle$ and $|\,1\rangle$. Since as we have seen in section 4.3 (page 117)

$$|\!\rightarrow\rangle \;=\; \frac{1}{\sqrt{2}}\left(|\,0\rangle + |\,1\rangle\right)$$

$$|\!\leftarrow\rangle \;=\; \frac{1}{\sqrt{2}}\left(|\,0\rangle - |\,1\rangle\right)$$

$|\!\rightarrow\rangle$ and $|\!\leftarrow\rangle$ are invariants of quantum NOT.

This implies that both $|\!\rightarrow\rangle$ and $|\!\leftarrow\rangle$ are invariants of quantum NOT:

$$\neg \,|\!\rightarrow\rangle = \neg \left(\frac{1}{\sqrt{2}}\left(|\,0\rangle + |\,1\rangle\right)\right) = \frac{1}{\sqrt{2}}\left(\neg \,|\,0\rangle + \neg \,|\,1\rangle\right)$$

$$= \frac{1}{\sqrt{2}}\left(|\,1\rangle + |\,0\rangle\right) = |\!\rightarrow\rangle$$

and

$$\neg \,|\!\leftarrow\rangle = \neg \left(\frac{1}{\sqrt{2}}\left(|\,0\rangle - |\,1\rangle\right)\right) = \frac{1}{\sqrt{2}}\left(\neg \,|\,0\rangle - \neg \,|\,1\rangle\right)$$

$$= \frac{1}{\sqrt{2}}\left(|\,1\rangle - |\,0\rangle\right) = -\,|\!\leftarrow\rangle$$

The minus sign in front of $|\!\leftarrow\rangle$ vanishes when we switch from the basis description of qubit states to fiducial vectors, so we can ignore it here. In other words, we can say that quantum NOT leaves $\boldsymbol{e}_x$ and $-\boldsymbol{e}_x$ unchanged.

Quantum NOT *swaps $\boldsymbol{e}_y$ and $-\boldsymbol{e}_y$.*

Similarly, it is easy to see that

$$| \otimes \rangle \;=\; \frac{1}{\sqrt{2}} \left(| \, 0 \, \rangle + i \, | \, 1 \, \rangle \right)$$

$$| \odot \rangle \;=\; \frac{1}{\sqrt{2}} \left(| \, 0 \, \rangle - i \, | \, 1 \, \rangle \right)$$

which implies that

$$\neg \, | \otimes \rangle \;=\; i \, | \odot \rangle$$

$$\neg \, | \odot \rangle \;=\; -i \, | \otimes \rangle$$

We can again ignore the factors i and $-i$ that appear in front of $| \odot \rangle$ and $| \otimes \rangle$ on the right hand side of the equations above, because they vanish when we switch to the fiducial vector formalism. What the above says, in effect, is that quantum NOT switches $\boldsymbol{e}_y$ to $-\boldsymbol{e}_y$ and vice versa.

These two additional observations define quantum NOT uniquely as the rotation of the Bloch ball by $180°$ about $\boldsymbol{e}_x$ and not some other axis.

Quantum NOT *is the rotation by* $180°$ *about* $\boldsymbol{e}_x$.

Having narrowed the definition and implementation of quantum NOT so, we can define another quantum gate, which is called *the square root of* NOT. If instead of rotating the Bloch ball by $180°$ about $\boldsymbol{e}_x$ we were to rotate it by $90°$ only, we would have to follow this operation with another rotation by $90°$ in order to complete the NOT. This is shown in figure 4.3.
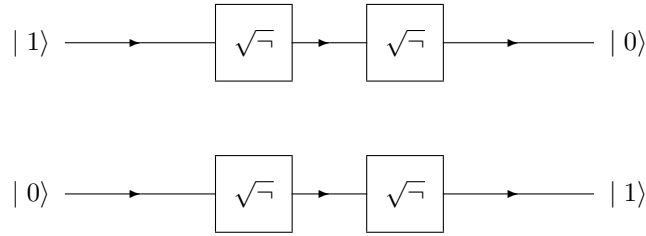
Square root of NOT

Figure 4.3: The square root of NOT.

The corresponding arithmetic definition of the square root of NOT is

$$\sqrt{\neg}\sqrt{\neg} \, | \, 0 \, \rangle \;=\; | \, 1 \, \rangle$$

$$\sqrt{\neg}\sqrt{\neg} \, | \, 1 \, \rangle \;=\; | \, 0 \, \rangle$$

Now, if we can have the square root of NOT, could we have a non-trivial square root of the identity? Such an operation is shown in figure 4.4. It is called the Hadamard rotation after the French mathematician Jacques-Salomon Hadamard (1865-1963).

Hadamard rotation, $\boldsymbol{HH} = \boldsymbol{1}$

The Hadamard rotation $\boldsymbol{H}$ rotates the Bloch ball about the direction that bisects the right angle between $-\boldsymbol{e}_x$ and $\boldsymbol{e}_z$, as shown in figure 4.4, by 180°. Because the direction of the axis of the Hadamard rotation is inclined by 45° with respect to $-\boldsymbol{e}_x$ and $\boldsymbol{e}_z$, the rotation swaps $-\boldsymbol{e}_x$ and $\boldsymbol{e}_z$. The Hadamard rotation also swaps $-\boldsymbol{e}_z$ and $\boldsymbol{e}_x$. Repeating it twice restores the Bloch Ball to its original orientation, hence $\boldsymbol{HH} = \mathbf{1}$.
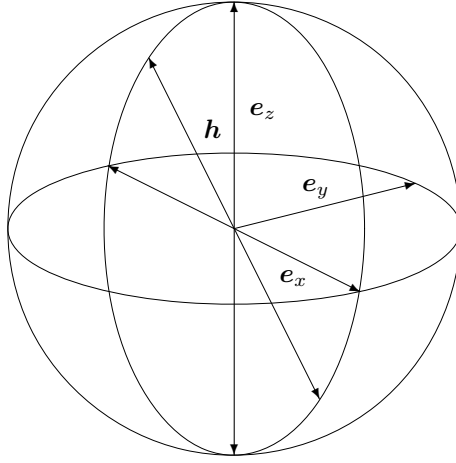
Figure 4.4: The Hadamard rotation $\boldsymbol{H}$ rotates the Bloch ball about vector $\boldsymbol{h}$, which bisects the right angle between $-\boldsymbol{e}_x$ and $\boldsymbol{e}_z$, by 180°. This swaps $-\boldsymbol{e}_z$ and $\boldsymbol{e}_x$. At the same time, $\boldsymbol{e}_z$ and $-\boldsymbol{e}_x$ are swapped too. Applying $\boldsymbol{H}$ twice results in the rotation by 360° about $\boldsymbol{h}$, which brings the Bloch ball to its original orientation.

We can define the Hadamard rotation diagrammatically as shown in figure 4.5. The arithmetic definition of the Hadamard rotation is as follows:

$$\boldsymbol{H} \, | \, 0 \rangle \;\; = \;\; \frac{1}{\sqrt{2}} \left(| \, 0 \rangle + | \, 1 \rangle \right)$$

$$\boldsymbol{H} \, | \, 1 \rangle \;\; = \;\; \frac{1}{\sqrt{2}} \left(| \, 0 \rangle - | \, 1 \rangle \right)$$

It is easy to see both from figure 4.4 and from the following calculation that the Hadamard rotation swaps $\boldsymbol{e}_y$ and $-\boldsymbol{e}_y$:

$$\boldsymbol{H} \, | \otimes \rangle = \boldsymbol{H} \frac{1}{\sqrt{2}} \left(| \, 0 \rangle + i \, | \, 1 \rangle \right)$$

$$= \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} \left(| \, 0 \rangle + | \, 1 \rangle \right) + \frac{i}{\sqrt{2}} \left(| \, 0 \rangle - | \, 1 \rangle \right) \right)$$

$$| 1\rangle \longrightarrow \boxed{\boldsymbol{H}} \longrightarrow \quad |\leftarrow\rangle = \tfrac{1}{\sqrt{2}}\left(| 0\rangle - | 1\rangle\right)$$

$$| 0\rangle \longrightarrow \boxed{\boldsymbol{H}} \longrightarrow \quad |\rightarrow\rangle = \tfrac{1}{\sqrt{2}}\left(| 0\rangle + | 1\rangle\right)$$
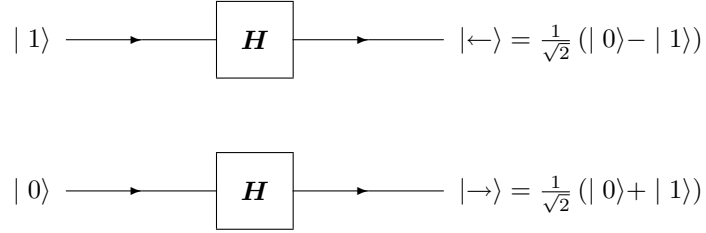
Figure 4.5: Diagrammatic representation of the Hadamard gate.

$$= \frac{1}{\sqrt{2}}\left(\frac{1+i}{\sqrt{2}}\,| 0\rangle + \frac{1-i}{\sqrt{2}}\,| 1\rangle\right) = \frac{1+i}{\sqrt{2}}\,\frac{1}{\sqrt{2}}\left(| 0\rangle - i\,| 1\rangle\right)$$
$$= \frac{1+i}{\sqrt{2}}\,| \odot\rangle$$

As before we can ignore the factor $(1 + i)/\sqrt{2}$, because it vanishes in the translation between the basis states and fiducial vector descriptions of the qubit.

Operations such as the Hadamard rotation and the square root of NOT are a *Similarity to* reflection of the fact that by using the controlled Larmor precessions about various *analog* directions in space, or a combination of the Larmor precession and the Rabi oscilla- *computing* tions we can do with a qubit all that we can do with a normal ball. We can rotate its Bloch ball in any way we wish thus implementing an arbitrary continuous mapping between any two points on the Bloch sphere. By adding various dissipative operations such as the measurement we can dig *into* the interior of the Bloch ball too.

This makes the qubit a very rich object, markedly richer than its classical cousin, the bit, which can only assume one of two discrete values. These riches resemble analog, or fuzzy logic computing, where a computational element can assume any real value within a certain range. Indeed, a classical ball, e.g., a ping-pong ball, provides us with the computational equivalent of the qubit's Bloch sphere.

The only snag is our inability to extract the full information about r from a *Qubit states can* single measurement. In order to ascertain the state of the qubit we have to perform *be ascertained* a very large number of measurements on it, in other words, we have to fully explore *only by* its statistical ensemble. *exploring their statistical*

There are two fundamental ways to do this. *ensembles.*

The first way is to work with a single qubit *sequentially*, as we have seen in the *Sequential* quantronium example. We would prepare the qubit in some well defined initial *exploration of the statistical ensemble*

state. Then we would perform some operations on the qubit and finally we would send the qubit through the beam splitter in order to see if it emerges in state $|\uparrow\rangle$ or $|\downarrow\rangle$. We would have to perform this procedure, for example 50,000 times, as was the case in the quantronium example, in order go gather sufficient statistics that would give us r^z. We would then have to repeat all these operations without change but on the output we would modify the beam splitting chamber so that the qubit would emerge from it in state $|\rightarrow\rangle$ or $|\leftarrow\rangle$. After some 50,000 of such measurements we would get a fair idea about r^x. Then we would have to repeat this whole procedure once more, but this time we would rotate the beam splitting apparatus so that the qubit would emerge from it in state $|\otimes\rangle$ or $|\odot\rangle$, which, after sufficient statistics had been collected, would yield r^y.

Parallel exploration of the statistical ensemble

The other way to approach the measurement of $\boldsymbol{r}$ is to work with millions, perhaps even billions of identically prepared qubits and send them in the form of a particle beam through various gates all at the same time. It will be then enough to rotate the beam splitting apparatus in various ways and to measure output beam intensities in order to reconstruct $\boldsymbol{r}$. This is a quick and very accurate way of doing things, because the large number of qubits ensures that we can average random errors away and obtain very precise distributions. The underlying assumption here is that the qubits do not interact with one another, or, if they do, that such interaction can be averaged away too.

4.10   Taking qubits for a ride

So far our qubits were stationary. A qubit would hang somewhere suspended either in a solution (this is how nuclear magnetic resonance experiments are carried out) or drifting in vacuum (this would be a qubit trapped in a potential well, or a qubit in a particle beam), or printed on a circuit board, and we would subject it to either static or buzzing "magnetic field" in order to manipulate its quantum state.

Manipulating qubits with rapid magnetic field sweeps

But there are also other ways to manipulate qubits. For example, one may subject a qubit to a sweeping magnetic field. Such magnetic field sweeps can be more precise and much faster than manipulating qubits gently by buzzing them with *weak* magnetic oscillations. And speed matters for many reasons. Faster gates mean faster computers for starters. But faster gates also mean that quantum information can be processed *before* the qubits that the information is encoded on decohere. Because of qubits' great sensitivity to the environment, quantum computing is always a race against time.

But in this section we are going to look at a yet another way of manipulating qubits that is in itself very interesting and that may play a role in the future of

quantum computing, although there hasn't been much activity in this field yet. We are going to take qubits for a ride. A slow ride in a parameter space.

> In the process we are also going to exercise all that we have learnt in this chapter: various formulas of the unitary formalism, whatever we learnt about the Hamiltonian and its eigenstates, the Schrödinger equation, etc.

How would a qubit's state change if we were to drag it from A to B along a certain trajectory through space filled with a "magnetic" field $\boldsymbol{B}$ that varies from location to location?

This question proved remarkably fruitful and it was only answered by Sir Michael V. Berry of Bristol University in a paper he published in the Proceedings of the Royal Society of London in 1984 [6]. For this achievement Berry was awarded a prestigious Wolf Prize in 1998 together with Yakir Aharonov, who discovered a related Aharonov-Bohm effect in 1959[5]. But the Aharonov-Bohm effect was specific to charged particles in the presence of the magnetic potential, whereas Berry effect applies to every quantum system, electromagnetic or not. It applies to systems that are not necessarily qubits too.

In practice moving a qubit physically from A to B is a difficult, if not impossible, endeavour. Qubits are extremely delicate, so there is no way to move them, even touch them, without destroying their quantum state at the same time. But since motion is relative, we can always move the environment around the qubit instead... and this is how the related experiments are carried out usually.

Moving the environment around a qubit is equivalent to moving the qubit itself.

4.10.1 Dragging a qubit along an arbitrary trajectory

Our starting point is the Schrödinger equation (4.41):

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t} \mid \Psi(t)\rangle = \boldsymbol{H}(t) \mid \Psi(t)\rangle$$

The environment the qubit is immersed in is represented by the Hamiltonian $\boldsymbol{H}$. As we move the qubit around a trajectory given by $\boldsymbol{x}(t)$, where $\boldsymbol{x} = (x, y, z)$ is the *guiding vector*, its environment changes and this can be described by making the Hamiltonian an explicit function of position $\boldsymbol{x}(t)$. Once the Hamiltonian is an explicit function of position, so must be the state vector, and so we end up with

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t} \mid \Psi(\boldsymbol{x}(t))\rangle = \boldsymbol{H}\left(\boldsymbol{x}(t)\right) \mid \Psi(\boldsymbol{x}(t))\rangle$$

---

[5]The reason why David Bohm was not awarded the prize was because he was dead by then. Bohm died of heart attack in London in 1992.

At every point along the trajectory the Hamiltonian matrix has some eigenvectors, two eigenvectors in the case of a qubit. They will vary from a point to a point changing their direction, though not length, because this is all done within the unitary formalism. Let us call these position dependent eigenvectors $\mid n(\boldsymbol{x}(t))\rangle$ so that

$$\boldsymbol{H}(\boldsymbol{x}(t)) \mid n(\boldsymbol{x}(t))\rangle = E_n(\boldsymbol{x}(t)) \mid n(\boldsymbol{x}(t))\rangle$$

where $E_n(\boldsymbol{x}(t))$ is the eigenvalue of $\boldsymbol{H}(\boldsymbol{x}(t))$ that corresponds to $\mid n(\boldsymbol{x}(t))\rangle$ at $\boldsymbol{x}(t)$.

Adiabatic evolution

Now we are going to introduce an important concept of an *adiabatic* motion. Suppose that a qubit is in an eigenstate $\mid n(\boldsymbol{x}(0))\rangle$ at the beginning. As we move the qubit ever so gently and *slowly* around $\boldsymbol{x}(t)$ the qubit has enough time to "thermalize" at every $\boldsymbol{x}(t)$, i.e., to adjust itself to the local Hamiltonian at this position, so that it remains in an eigenstate, even though the eigenstate itself changes.

Solution to the adiabatic motion problem postulated

Inspired by equation (4.46) we shall seek a solution to our adiabatic qubit transfer problem in the form

$$\mid \Psi(t)\rangle = e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\, \mathrm{d}t'\right)/\hbar} e^{i\gamma_n(t)} \mid n(\boldsymbol{x}(t))\rangle \tag{4.52}$$

We have the exponential with the time integral of the eigenvalue here, but we also allow for an additional time-dependent phase factor $e^{i\gamma_n(t)}$.

Now we plug this solution into the Schrödinger equation and the equation should tell us something about the way the gamma factor, $\gamma_n(t)$ relates to the eigenstates and, possibly, to the eigenenergies as well.

The right hand side of the Schrödinger equation, $\boldsymbol{H} \mid \Psi\rangle$ is easy. Here it simply translates into

$$E_n(\boldsymbol{x}(t)) \mid \Psi(t)\rangle$$

The left hand side, i.e., the time-derivative side of the equation, is somewhat more problematic, because the proposed solution (4.52) depends on time in a rather complicated way. But, at the end of the day there are just three factors here:

$$\mid \Psi(t)\rangle = e^{\mathrm{integral}} e^{\mathrm{gamma}} \mid \mathrm{eigenvector}\rangle$$

so

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \mid \Psi(t)\rangle \;=\; & \left(\frac{\mathrm{d}}{\mathrm{d}t} e^{\mathrm{integral}}\right) e^{\mathrm{gamma}} \mid \mathrm{eigenvector}\rangle \\
& + e^{\mathrm{integral}} \left(\frac{\mathrm{d}}{\mathrm{d}t} e^{\mathrm{gamma}}\right) \mid \mathrm{eigenvector}\rangle \\
& + e^{\mathrm{integral}} e^{\mathrm{gamma}} \left(\frac{\mathrm{d}}{\mathrm{d}t} \mid \mathrm{eigenvector}\rangle\right)
\end{aligned}
$$

The easiest of the three time derivatives is the derivative of the gamma factor. Here it is just:

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{i\gamma_n(t)} = i e^{i\gamma_n(t)} \frac{\mathrm{d}}{\mathrm{d}t} \gamma_n(t)$$

The other exponential, the one with the integral in it, throws out $-i/\hbar$ and the exponential itself and then we have to find a time derivative of the integral. But the time derivative of the integral is simply the integrated function $\boldsymbol{E_n}$ and so:

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\, \mathrm{d}t'\right)/\hbar} = -\frac{i}{\hbar} e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\, \mathrm{d}t'\right)/\hbar} E_n(\boldsymbol{x}(t))$$

Finally let us have a look at the time derivative of the eigenstate $\mid n(\boldsymbol{x}(t))\rangle$. The eigenstate is a Hilbert-vector valued function of position, which then itself is a function of time. The time derivative of $\mid n\rangle$ is therefore

$$\frac{\partial n(\boldsymbol{x}(t))}{\partial \boldsymbol{x}} \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}$$

The expression $\partial/\partial \boldsymbol{x}$ is an exotic way of writing a gradient $\boldsymbol{\nabla}$, so we can rewrite the above as

$$\frac{\mathrm{d}}{\mathrm{d}t} \mid n(\boldsymbol{x}(t))\rangle = \mid \boldsymbol{\nabla} n(\boldsymbol{x}(t))\rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}$$

Before we go any further, let me explain this expression, $\mid \boldsymbol{\nabla} n(\boldsymbol{x}(t))\rangle \cdot \mathrm{d}\boldsymbol{x}(t)/\mathrm{d}t$ some more. The reason this expression requires explaining is because we are mixing here vector objects that belong to very different spaces. There are normal 3-dimensional geometric space vectors and vector operators in it, namely $\boldsymbol{\nabla}$ and $\boldsymbol{x}$ and then we have a Hilbert space vector $\mid n\rangle$ in it too.

Any spinor, including $\mid n\rangle$, can be decomposed into basis spinors, e.g., $\mid\uparrow\rangle$ and $\mid\downarrow\rangle$

$$
\begin{aligned}
\mid n\rangle &= n_\uparrow \mid\uparrow\rangle + n_\downarrow \mid\downarrow\rangle \\
&= \sum_{m=\uparrow,\downarrow} n^m \mid m\rangle \\
&= \sum_{m=\uparrow,\downarrow} \mid m\rangle\langle m \mid n\rangle
\end{aligned}
$$

Here n^m does not mean n to the power of m. It means the m-th component of n, and it is, as we have seen in section 4.4, equations (4.25) and (4.26), simply $\langle m \mid n(\boldsymbol{x})\rangle$, which is a normal complex valued function of position $\boldsymbol{x}$, a function that can be differentiated.

Taking a 3-D space gradient of $\mid n\rangle$ means the following

$$\mid \boldsymbol{\nabla} n(\boldsymbol{x})\rangle = \sum_{m=\uparrow,\downarrow} \boldsymbol{\nabla} n^m(\boldsymbol{x}) \mid m\rangle$$

$$
\begin{aligned}
&= \sum_{m=\uparrow,\downarrow} \sum_{i=x,y,z} \frac{\partial n^m(\boldsymbol{x})}{\partial x^i} \boldsymbol{e}_i \otimes \mid m \rangle \\
&= \sum_{m=\uparrow,\downarrow} \sum_{i=x,y,z} \boldsymbol{e}_i \otimes \mid m \rangle \frac{\partial \langle m \mid n(\boldsymbol{x}) \rangle}{\partial x^i}
\end{aligned}
$$

This is a tensor product with one leg, $\boldsymbol{e}_i$, standing in the 3-D space and the other leg, $\mid m \rangle$, standing in the spinor space. Now we are going to take a 3-D space scalar (dot) product of this with

$$
\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \sum_{j=x,y,z} \frac{\mathrm{d}x^j(t)}{\mathrm{d}t} \boldsymbol{e}_j
$$

Of course, we cannot contract $\boldsymbol{e}_j$ with $\mid m \rangle$. We can only contract it with $\boldsymbol{e}_i$, and so we end up with

$$
\mid \boldsymbol{\nabla} n(\boldsymbol{x}) \rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \sum_{m=\uparrow,\downarrow} \mid m \rangle \left(\sum_{i=x,y,z} \frac{\partial \langle m \mid n(\boldsymbol{x}) \rangle}{\partial x^i} \frac{\mathrm{d}x^i(t)}{\mathrm{d}t} \right)
$$

The 3-D vectors of this expression eat each other in the frenzy of a dot product and leave a 3-D scalar behind. This scalar is

$$
\sum_{i=x,y,z} \frac{\partial \langle m \mid n(\boldsymbol{x}) \rangle}{\partial x^i} \frac{\mathrm{d}x^i(t)}{\mathrm{d}t}
$$

and it is this scalar that is now used as the coefficient in the spinor's expansion into the basis spinors of the Hilbert space.

Now we have to put it all together into $i\hbar \mathrm{d} \mid \Psi \rangle / \mathrm{d}t = \boldsymbol{H} \mid \Psi \rangle$:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \mid \Psi(t) \rangle &= \frac{\mathrm{d}}{\mathrm{d}t} \left(e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar} e^{i\gamma_n(t)} \mid n(\boldsymbol{x}(t)) \rangle \right) \\
&= -\frac{i}{\hbar} e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar} E_n(\boldsymbol{x}(t)) e^{i\gamma_n(t)} \mid n(\boldsymbol{x}(t)) \rangle \\
&\quad + e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar} i e^{i\gamma_n(t)} \frac{\mathrm{d}}{\mathrm{d}t} \gamma_n(t) \mid n(\boldsymbol{x}(t)) \rangle \\
&\quad + e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar} e^{i\gamma_n(t)} \mid \boldsymbol{\nabla} n(\boldsymbol{x}(t)) \rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} \\
&= -\frac{i}{\hbar} E_n(\boldsymbol{x}(t)) \mid \Psi(t) \rangle + i \frac{\mathrm{d}\gamma_n(t)}{\mathrm{d}t} \mid \Psi(t) \rangle \\
&\quad + e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar} e^{i\gamma_n(t)} \mid \boldsymbol{\nabla} n(\boldsymbol{x}(t)) \rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}
\end{aligned}
$$

We are nearly there. Let us multiply this by $i\hbar$ and then let us match it against $E_n \mid \Psi\rangle$ on the right hand side:

$$E_n(\boldsymbol{x}(t)) \mid \Psi(t)\rangle - \hbar\frac{\mathrm{d}\gamma_n(t)}{\mathrm{d}t} \mid \Psi(t)\rangle$$

$$+ i\hbar e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar}e^{i\gamma_n(t)} \mid \boldsymbol{\nabla}n(\boldsymbol{x}(t))\rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}$$

$$= E_n(\boldsymbol{x}(t)) \mid \Psi(t)\rangle$$

We can immediately see $E_n \mid \Psi\rangle$ on both sides of the equation. We can also see a lot of other stuff on the left hand side, which we don't want. With this other stuff out of the way we would get $E_n \mid \Psi\rangle = E_n \mid \Psi\rangle$, which is a perfectly fine way of making the Schrödinger equation happy.

So we arrive at the condition: "this other stuff ought to vanish". We translate this into the following equation:

$$-\hbar\frac{\mathrm{d}\gamma_n(t)}{\mathrm{d}t} \mid \Psi(t)\rangle + i\hbar e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar}e^{i\gamma_n(t)} \mid \boldsymbol{\nabla}n(\boldsymbol{x}(t))\rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = 0$$

But recall that $\mid \Psi\rangle$ itself contains three terms and it is good to write them explicitly here:

$$-\hbar\frac{\mathrm{d}\gamma_n(t)}{\mathrm{d}t}e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar}e^{i\gamma_n(t)} \mid n(\boldsymbol{x}(t))\rangle$$

$$+ i\hbar e^{-i\left(\int_0^t E_n(\boldsymbol{x}(t'))\,\mathrm{d}t'\right)/\hbar}e^{i\gamma_n(t)} \mid \boldsymbol{\nabla}n(\boldsymbol{x}(t))\rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}$$

$$= 0$$

Now we can throw a lot of stuff away. First we divide both sides by the exponentials, then we divide both sides by $\hbar$ and this leaves us with:

$$\frac{\mathrm{d}\gamma_n(t)}{\mathrm{d}t} \mid n(\boldsymbol{x}(t))\rangle = i \mid \boldsymbol{\nabla}n(\boldsymbol{x}(t))\rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}$$

Finally we multiply both sides by $\langle n(\boldsymbol{x}(t))\mid$ from the left. This eats $\mid n(\boldsymbol{x}(t))\rangle$ on the left hand side leaving a pure time derivative of γ_n there:

$$\frac{\mathrm{d}\gamma_n(t)}{\mathrm{d}t} = i\langle n(\boldsymbol{x}(t)) \mid \boldsymbol{\nabla}n(\boldsymbol{x}(t))\rangle \cdot \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} \tag{4.53}$$

This equation has an obvious solution in the form of a line integral

$$\gamma_n(C) = i\int_C \langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla}n(\boldsymbol{x})\rangle \cdot \mathrm{d}\boldsymbol{x} \tag{4.54}$$

The differential equation for the geometric phase in adiabatic motion

which says that as we move a qubit along trajectory C, its phase γ_n accumulates contributions of $\langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla} n(\boldsymbol{x})\rangle \cdot \Delta\boldsymbol{x}$ along the line. It is important to notice that this accumulation does not depend on how fast or how slow we move the qubit. It depends only on the variation of $|n(\boldsymbol{x})\rangle$ along the trajectory C.

The integral is purely imaginary, therefore $\gamma_n(C)$ is purely real. This is important, because if $\gamma_n(C)$ had an imaginary component the exponential $e^{i\gamma_n}$ would change the length of vector $|\Psi\rangle$, and this cannot happen within the confines of the unitary formalism.

That the integral is purely imaginary can be seen as follows:

$$
\begin{aligned}
0 &= \boldsymbol{\nabla}1 = \boldsymbol{\nabla}\langle n \mid n\rangle \\
&= \langle \boldsymbol{\nabla} n \mid n\rangle + \langle n \mid \boldsymbol{\nabla} n\rangle \\
&= \langle n \mid \boldsymbol{\nabla} n\rangle^* + \langle n \mid \boldsymbol{\nabla} n\rangle \\
&= 2\Re\langle n \mid \boldsymbol{\nabla} n\rangle
\end{aligned}
$$

4.10.2   A closed trajectory case

Equation (4.54) can be transformed farther if the trajectory C encloses a surface S so that

$$C = \partial S$$

where ∂S means "the edge of S". In this case $\int_C$ becomes $\oint_{\partial S}$ and we can invoke the Stokes theorem that converts a line integral over the edge of a surface into a *curl* integral over the surface itself:

$$
\oint_{\partial S} \langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla} n(\boldsymbol{x})\rangle \cdot \mathrm{d}\boldsymbol{x} = \int_S \left(\boldsymbol{\nabla} \times \langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla} n(\boldsymbol{x})\rangle\right) \cdot \mathrm{d}^2\boldsymbol{S}
$$

We should again stop here and explain this expression in terms of actual functions, vector and spinor components. What is being differentiated here and what is being "cross-producted"? $\langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla} n(\boldsymbol{x})\rangle$ is a 3-D vector field on the normal 3-D space. This vector field arises in the following way:

$$
\langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla} n(\boldsymbol{x})\rangle = \sum_{m=\uparrow,\downarrow} \sum_{i=x,y,z} n_m^*(\boldsymbol{x}) \left(\frac{\partial}{\partial x^i} n^m(\boldsymbol{x})\right) \boldsymbol{e}_i
$$

where $n_m^* = \langle n \mid m\rangle$ and $n^m = \langle m \mid n\rangle$. The spinor index m is summed away (we call such an index *saturated*) and we are left with just a 3-D space index i. Every i

term is then multiplied by $\boldsymbol{e}_i$ so that a vector field comes out. Now we act on this field with the *curl* operator $\boldsymbol{\nabla}\times$. The result is:

$$
\boldsymbol{\nabla} \times \langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla} n(\boldsymbol{x}) \rangle = \sum_{\substack{i \,=\, x,y,z \\ j \,=\, x,y,z \\ k \,=\, x,y,z}} \epsilon_{ijk} \boldsymbol{e}_i \frac{\partial}{\partial x^j} \left(\sum_{m=\uparrow,\downarrow} n_m^*(\boldsymbol{x}) \frac{\partial}{\partial x^k} n^m(\boldsymbol{x}) \right)
$$

where ϵ_{ijk} is the fully antisymmetric 3-D symbol. The spinor index m in this expression is saturated as before. So are the j and k 3-D space indexes, but this time they are saturated in the cross-product way so that, for example, the x-component of $\boldsymbol{\nabla} \times \langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla} n(\boldsymbol{x}) \rangle$ is

$$
\frac{\partial}{\partial y} \left(\sum_{m=\uparrow,\downarrow} n_m^*(\boldsymbol{x}) \frac{\partial}{\partial z} n^m(\boldsymbol{x}) \right) - \frac{\partial}{\partial z} \left(\sum_{m=\uparrow,\downarrow} n_m^*(\boldsymbol{x}) \frac{\partial}{\partial y} n^m(\boldsymbol{x}) \right)
$$

These expressions may look somewhat tedious, but they should not look scary. They are easy to understand in terms of what is what. And we are going to go some way still towards making them more usable.

The first thing to observe about

$$
\sum_{\substack{i \,=\, x,y,z \\ j \,=\, x,y,z \\ k \,=\, x,y,z}} \epsilon_{ijk} \boldsymbol{e}_i \frac{\partial}{\partial x^j} \left(\sum_{m=\uparrow,\downarrow} n_m^*(\boldsymbol{x}) \frac{\partial}{\partial x^k} n^m(\boldsymbol{x}) \right)
$$

is that the derivative $\partial/\partial x^j$ is going to hit $n_m^*(\boldsymbol{x})$ first, but then when it gets to $\partial n^m(\boldsymbol{x})/\partial x^k$ it'll give us zero, because ϵ_{ijk} is anti-symmetric in j and k, but $\frac{\partial}{\partial x^j}\frac{\partial}{\partial x^k}$ is symmetric in j and k. This is a yet another formulation of the rule that a *curl* of a gradient is zero, $\boldsymbol{\nabla} \times \boldsymbol{\nabla} n^m(\boldsymbol{x}) = 0$.

Consequently we are going to have that

$$
\boldsymbol{\nabla} \times \langle n(\boldsymbol{x}) \mid \boldsymbol{\nabla} n(\boldsymbol{x}) \rangle \;\; = \;\; \sum_{m=\uparrow,\downarrow} \sum_{\substack{i \,=\, x,y,z \\ j \,=\, x,y,z \\ k \,=\, x,y,z}} \epsilon_{ijk} \left(\frac{\partial n_m^*(\boldsymbol{x})}{\partial x^j} \right) \left(\frac{\partial n^m(\boldsymbol{x})}{\partial x^k} \right) \boldsymbol{e}_i
$$

Observe that $\langle \boldsymbol{\nabla} n \mid$ is not the same as $\mid \boldsymbol{\nabla} n \rangle$. On the index level one is n_m^* and the other one is n^m. This is just as well, because if they were the same then we

would have something like $(\boldsymbol{\nabla} n) \times (\boldsymbol{\nabla} n) = 0$ and the whole computation would be over.

Now, $\boldsymbol{\nabla} n^m$ is simply $\langle m \mid \boldsymbol{\nabla} n \rangle$ and $\boldsymbol{\nabla} n_m^*$ is $\langle \boldsymbol{\nabla} n \mid m \rangle$. Therefore we can rewrite the expression for γ_n as follows:

$$\gamma_n(\partial S) = i \int_S \sum_{m=\uparrow,\downarrow} \langle \boldsymbol{\nabla} n(\boldsymbol{x}) \mid m \rangle \times \langle m \mid \boldsymbol{\nabla} n(\boldsymbol{x}) \rangle \cdot \mathrm{d}^2 \boldsymbol{S}$$

because $\sum_{ijk} \epsilon_{ijk} \boldsymbol{e}_i u^j v^k = \boldsymbol{u} \times \boldsymbol{v}$.

This does not look simpler or more useful than equation (4.54), but we are now going to invoke two important facts.

The first fact is that $\mid n \rangle$ is an eigenvector of a local $\boldsymbol{H}$ at every point on surface S and therefore it is one of $\mid \uparrow \rangle$ or $\mid \downarrow \rangle$ at that point. Observe that since $\boldsymbol{\nabla} \langle n \mid n \rangle = 0$ we have that $\langle \boldsymbol{\nabla} n \mid n \rangle = -\langle n \mid \boldsymbol{\nabla} n \rangle$, therefore

$$\langle \boldsymbol{\nabla} n \mid n \rangle \times \langle n \mid \boldsymbol{\nabla} n \rangle = 0$$

because for any vector $\boldsymbol{v}$ we have that $\boldsymbol{v} \times (-\boldsymbol{v}) = 0$. For this reason we can rewrite our expression for γ_n yet again, as follows:

$$\gamma_n(\partial S) = i \int_S \sum_{m \neq n} \langle \boldsymbol{\nabla} n(\boldsymbol{x}) \mid m \rangle \times \langle m \mid \boldsymbol{\nabla} n(\boldsymbol{x}) \rangle \cdot \mathrm{d}^2 \boldsymbol{S}$$

where $\mid m \rangle$ is *the other* eigenvector. Even though for two dimensional systems such as qubits, the sum in this equation reduces to just one component, we are going to keep it, because this way the expression is going to be valid for systems with larger number of dimensions too – in this case $\mid m \rangle$ stands for *all the other* eigenvectors of the local Hamiltonian.

The second fact is that $\mid n \rangle$ must satisfy

$$\boldsymbol{H} \mid n \rangle = E_n \mid n \rangle$$

Let us apply the Nabla operator to both sides:

$$\boldsymbol{\nabla} (\boldsymbol{H} \mid n \rangle) = \boldsymbol{\nabla} (E_n \mid n \rangle)$$

This yields

$$(\boldsymbol{\nabla} \boldsymbol{H}) \mid n \rangle + \boldsymbol{H} \mid \boldsymbol{\nabla} n \rangle = (\boldsymbol{\nabla} E_n) \mid n \rangle + E_n \mid \boldsymbol{\nabla} n \rangle$$

Let us multiply this equation by $\langle m \mid$ from the left

$$\langle m \mid \boldsymbol{\nabla} \boldsymbol{H} \mid n \rangle + \langle m \mid \boldsymbol{H} \mid \boldsymbol{\nabla} n \rangle = \boldsymbol{\nabla} E_n \langle m \mid n \rangle + E_n \langle m \mid \boldsymbol{\nabla} n \rangle$$

Recall that $\langle m \mid \boldsymbol{H} = \langle m \mid E_m$. Also, since $\mid m \rangle \neq \mid n \rangle$ we have that $\langle m \mid n \rangle = 0$. In summary

$$\langle m \mid \boldsymbol{\nabla} \boldsymbol{H} \mid n \rangle + E_m \langle m \mid \boldsymbol{\nabla} n \rangle = E_n \langle m \mid \boldsymbol{\nabla} n \rangle$$

or

$$\langle m \mid \boldsymbol{\nabla} n \rangle = \frac{\langle m \mid \boldsymbol{\nabla} \boldsymbol{H} \mid n \rangle}{E_n - E_m}$$

A similar expression holds for $\langle \boldsymbol{\nabla} n \mid m \rangle$. This lets us rewrite our equation for $\gamma_n(\partial S)$ yet again:

$$\gamma_n(\partial S) = i \int_S \sum_{m \neq n} \frac{\langle n \mid \boldsymbol{\nabla} \boldsymbol{H} \mid m \rangle \times \langle m \mid \boldsymbol{\nabla} \boldsymbol{H} \mid n \rangle}{\left(E_n - E_m\right)^2} \cdot \mathrm{d}^2 \boldsymbol{S} \qquad (4.55)$$

Surface integral solution for the geometric phase arising from adiabatic motion around a closed loop

This equation should be read as follows. At every point of surface S we are going to differentiate the Hamiltoanian. This will produce *three* new operators in place of just one, they'll correspond to $\partial \boldsymbol{H}/\partial x$, $\partial \boldsymbol{H}/\partial y$ and $\partial \boldsymbol{H}/\partial z$. For a given state $\mid n \rangle$ and for each state $\mid m \rangle \neq \mid n \rangle$ (in case of a qubit there will be only one such state) we are going to have *three* numbers per point obtained by evaluating transition amplitudes $\langle n \mid \partial \boldsymbol{H}/\partial x \mid m \rangle$, $\langle n \mid \partial \boldsymbol{H}/\partial y \mid m \rangle$ and $\langle n \mid \partial \boldsymbol{H}/\partial z \mid m \rangle$ – thus forming a 3-D vector at this point. We are also going to have another three numbers obtained by evaluating $\langle m \mid \partial \boldsymbol{H}/\partial x \mid n \rangle$, $\langle m \mid \partial \boldsymbol{H}/\partial y \mid n \rangle$ and $\langle m \mid \partial \boldsymbol{H}/\partial z \mid n \rangle$ forming another 3-D vector at this point. We will have to take a cross product of these two vectors and divide it by $\left(E_n - E_m\right)^2$ then we'll have to take a scalar product of the resulting vector with the surface element $\mathrm{d}^2 \boldsymbol{S}$ and this will produce a ... number – just a normal complex number. For each point of surface S this operation has to be repeated for the remaining m-s, if there are such – this is only going to be the case in quantum systems with a larger number of dimensions than single qubits – and the results added. Finally, the operation has to be repeated for every other point of the surface, then the resulting numbers all summed up, the result multiplied by i and ... this is our γ_n.

A numerical procedure doing all this can be easily implemented and this means that we understand the formula. This is not always the case in quantum physics.

Equations (4.54) and (4.55) provide us with a method to evaluate the change of phase that accompanies the movement of a qubit (or any other quantum system, because we were sufficiently general here) along any trajectory and along a trajectory that encloses a surface respectively. The first equation (4.54) that looks quite simple is also very general. The second equation (4.55) looks more complicated, but this is because it is more specific.

4.10.3   A qubit in the rotating magnetic field

Although the title of this section "Taking qubits for a ride", as well as the wording that accompanies the derivation are highly suggestive of moving the qubits in the physical geometric 3-D space, the formulas derived and the reasoning itself are more general. They describe the movement of the qubit or any other quantum system in any parametric space. For example, if a qubit is subjected to slowly varying magnetic field, sufficiently slowly for the qubit's eigenstates to thermalize, i.e., align with the direction of the magnetic field, at all stages of the evolution – this can be also thought of as a "movement" of the qubit through ... the magnetic space. The only assumption we have made in deriving equation (4.55) was that the parametric space was 3-dimensional, so that both the Stokes formula and cross product manipulations could be applied. But no such assumption was made in deriving equation (4.54), which is therefore applicable to higher dimensional parametric spaces too.

Rotating magnetic field adiabatically around a qubit   Consider a situation in which a magnetic field vector $\boldsymbol{B}$ is rotated *adiabatically* around a qubit in a plane as shown in figure 4.6. The field's value does not change
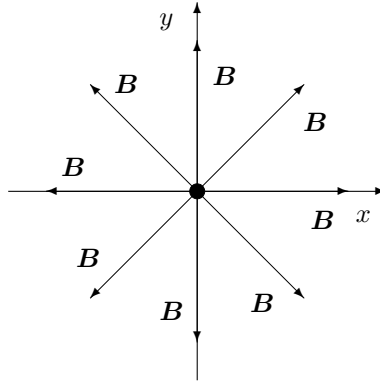
Figure 4.6: Magnetic field $\boldsymbol{B}$ rotates adiabatically around a qubit placed in the center of the figure.

throughout the rotation. This operation is equivalent to taking the qubit for a ride along a circle of radius B in the $\boldsymbol{B}$ space as shown in figure 4.7.

In order to evaluate a contribution that this operation is going to make to the phase $\gamma_n(\partial S)$ (remember that there is going to be a dynamic phase factor $\exp\left(-i \int_0^t E_n\left(\boldsymbol{x}(t')\right) \mathrm{d}t'\right)$ in the complete solution for $\mid \Psi(t)\rangle$ too) we need to use equation (4.55). But right here we have a conundrum, because the surface S in
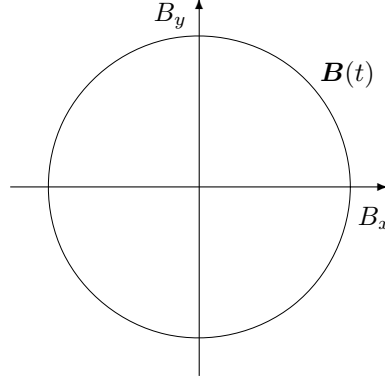
Figure 4.7: A qubit can be thought of as moving along the circle of radius B in the $\boldsymbol{B}$ space.

the $\boldsymbol{B}$ space, namely the circle of radius B, passes through $\boldsymbol{B} = 0$, where $E_\uparrow = E_\downarrow$ and equation (4.55) tells us that we should divide some such rather complicated expression made of a cross product and transition amplitudes by $(E_\uparrow - E_\downarrow)^2$. Alas, this expression is valid and the same for *any* surface S as long as its edge is the contour along which the qubit moves. So here we can use a different surface, for example, a hemi-sphere of radius B that stands on the great circle of the sphere, with the qubit moving along the perimeter of the great circle. The hemi-sphere stays away from $B = 0$ and so we don't have the problem. Furthermore, B is the same at every point of the hemi-sphere, so this will make our calculations easier.

Surface S may be deformed so as to keep it away from singularity. A hemi-sphere is a good surface to choose because B is constant on it and because we know its surface area.

We are going to evaluate $\gamma_n(\partial S)$ for $n =\uparrow, \downarrow$. The sum in equation (4.55) reduces to a single component only, because our qubit system is two dimensional. And so we have:

$$\gamma_\uparrow(\partial S) = i \int_S \frac{\langle\uparrow| \boldsymbol{\nabla H} |\downarrow\rangle \times \langle\downarrow| \boldsymbol{\nabla H} |\uparrow\rangle}{(E_\uparrow - E_\downarrow)^2} \cdot \mathrm{d}^2\boldsymbol{S}$$

$$\gamma_\downarrow(\partial S) = i \int_S \frac{\langle\downarrow| \boldsymbol{\nabla H} |\uparrow\rangle \times \langle\uparrow| \boldsymbol{\nabla H} |\downarrow\rangle}{(E_\downarrow - E_\uparrow)^2} \cdot \mathrm{d}^2\boldsymbol{S}$$

Now, let us have a look at $\boldsymbol{\nabla H}$. This is easy to evaluate because $\boldsymbol{H} = -\mu \left(B_x\boldsymbol{\sigma}_x + B_y\boldsymbol{\sigma}_y + B_z\boldsymbol{\sigma}_z\right)$■ and $\boldsymbol{\nabla} = (\partial/\partial B_x, \partial/\partial B_y, \partial/\partial B_z)$ and so

$$\boldsymbol{\nabla H} = -\mu \begin{pmatrix} \boldsymbol{\sigma}_x \\ \boldsymbol{\sigma}_y \\ \boldsymbol{\sigma}_z \end{pmatrix}$$

Consequently

$$
\gamma_\uparrow(\partial S) = i \int_S \frac{\mu^2}{(E_\uparrow - E_\downarrow)^2}
\begin{pmatrix}
\langle\uparrow|\,\boldsymbol{\sigma}_y\,|\downarrow\rangle\langle\downarrow|\,\boldsymbol{\sigma}_z\,|\uparrow\rangle - \langle\uparrow|\,\boldsymbol{\sigma}_z\,|\downarrow\rangle\langle\downarrow|\,\boldsymbol{\sigma}_y\,|\uparrow\rangle \\
\langle\uparrow|\,\boldsymbol{\sigma}_z\,|\downarrow\rangle\langle\downarrow|\,\boldsymbol{\sigma}_x\,|\uparrow\rangle - \langle\uparrow|\,\boldsymbol{\sigma}_x\,|\downarrow\rangle\langle\downarrow|\,\boldsymbol{\sigma}_z\,|\uparrow\rangle \\
\langle\uparrow|\,\boldsymbol{\sigma}_x\,|\downarrow\rangle\langle\downarrow|\,\boldsymbol{\sigma}_y\,|\uparrow\rangle - \langle\uparrow|\,\boldsymbol{\sigma}_y\,|\downarrow\rangle\langle\downarrow|\,\boldsymbol{\sigma}_x\,|\uparrow\rangle
\end{pmatrix} \cdot \mathrm{d}^2 \boldsymbol{S}
$$

and

$$
\gamma_\downarrow(\partial S) = i \int_S \frac{\mu^2}{(E_\uparrow - E_\downarrow)^2}
\begin{pmatrix}
\langle\downarrow|\,\boldsymbol{\sigma}_y\,|\uparrow\rangle\langle\uparrow|\,\boldsymbol{\sigma}_z\,|\downarrow\rangle - \langle\downarrow|\,\boldsymbol{\sigma}_z\,|\uparrow\rangle\langle\uparrow|\,\boldsymbol{\sigma}_y\,|\downarrow\rangle \\
\langle\downarrow|\,\boldsymbol{\sigma}_z\,|\uparrow\rangle\langle\uparrow|\,\boldsymbol{\sigma}_x\,|\downarrow\rangle - \langle\downarrow|\,\boldsymbol{\sigma}_x\,|\uparrow\rangle\langle\uparrow|\,\boldsymbol{\sigma}_z\,|\downarrow\rangle \\
\langle\downarrow|\,\boldsymbol{\sigma}_x\,|\uparrow\rangle\langle\uparrow|\,\boldsymbol{\sigma}_y\,|\downarrow\rangle - \langle\downarrow|\,\boldsymbol{\sigma}_y\,|\uparrow\rangle\langle\uparrow|\,\boldsymbol{\sigma}_x\,|\downarrow\rangle
\end{pmatrix} \cdot \mathrm{d}^2 \boldsymbol{S}
$$

Observe that $\gamma_\uparrow(\partial S) = -\gamma_\downarrow(\partial S)$. This is easy to see once the integrals have been written out in detail as above.

Remember that $|\uparrow\rangle$ and $|\downarrow\rangle$ are the eigenvectors of the local Hamiltonian and the local Hamiltonian, which is represented by vector $\boldsymbol{B}$ may point in any direction, not necessarily in the $\boldsymbol{e}_z$ direction. Consequently, $|\uparrow\rangle$ and $|\downarrow\rangle$ may not be equivalent to $\binom{1}{0}$ and $\binom{0}{1}$, which are the eigenvectors of $\boldsymbol{\sigma}_z$. Nevertheless we can evaluate $(E_\uparrow - E_\downarrow)^2$, which is $(2\mu B)^2$ and

$$
\frac{\mu^2}{(E_\uparrow - E_\downarrow)^2} = \frac{\mu^2}{4\mu^2 B^2} = \frac{1}{4B^2}
$$

We may represent $|\uparrow\rangle$ and $|\downarrow\rangle$ by $\binom{1}{0}$ and $\binom{0}{1}$ respectively at a selected point of surface S.
At a given point of surface S we *can* rotate our system of coordinates so that $\boldsymbol{H} = -\mu B \boldsymbol{\sigma}_z$ – just at this point only. This does not affect $\boldsymbol{\nabla} \boldsymbol{H}$, which still has three non-vanishing components, but it lets us represent $|\uparrow\rangle$ and $|\downarrow\rangle$ at this point and at this point only as $\binom{1}{0}$ and $\binom{0}{1}$. At this point then we can evaluate

$$
\begin{pmatrix}
(1,0)\boldsymbol{\sigma}_y \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1)\boldsymbol{\sigma}_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} - (1,0)\boldsymbol{\sigma}_z \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1)\boldsymbol{\sigma}_y \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
(1,0)\boldsymbol{\sigma}_z \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1)\boldsymbol{\sigma}_x \begin{pmatrix} 1 \\ 0 \end{pmatrix} - (1,0)\boldsymbol{\sigma}_x \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1)\boldsymbol{\sigma}_z \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
(1,0)\boldsymbol{\sigma}_x \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1)\boldsymbol{\sigma}_y \begin{pmatrix} 1 \\ 0 \end{pmatrix} - (1,0)\boldsymbol{\sigma}_y \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1)\boldsymbol{\sigma}_x \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\end{pmatrix}
$$

Before we go any further, recall that $\binom{1}{0}$ and $\binom{0}{1}$ are eigenvectors of $\boldsymbol{\sigma}_z$. For this reason all occurrences of $\binom{1}{0}\boldsymbol{\sigma}_z\binom{0}{1}$ and $\binom{0}{1}\boldsymbol{\sigma}_z\binom{1}{0}$ must vanish and by doing so they'll take down the first and the second row of the large, fat vector above with them. The only row that's going to survive is the third row. And this row is:

$$
(1,0)\boldsymbol{\sigma}_x \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1)\boldsymbol{\sigma}_y \begin{pmatrix} 1 \\ 0 \end{pmatrix} - (1,0)\boldsymbol{\sigma}_y \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1)\boldsymbol{\sigma}_x \begin{pmatrix} 1 \\ 0 \end{pmatrix}
$$

$$= (1,0) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$- (1,0) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0,1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$= i - (-i) = 2i$$

And so at this particular point the vector is

$$\begin{pmatrix} 0 \\ 0 \\ 2i \end{pmatrix} = 2i \frac{\boldsymbol{B}}{B}$$

By writing it in this form $\boldsymbol{B}/B$ we make it independent of the system of coordinates we have chosen to evaluate this expression, so that we can use it elsewhere too.

Now we can finally wrap it all together and this is the expression we get for $\gamma_\uparrow(\partial S)$ (and for $\gamma_\downarrow = -\gamma_\uparrow$ as well):

$$\gamma_\uparrow(\partial S) = i \int_S \frac{1}{4B^2} \frac{2i\boldsymbol{B}}{B} \cdot \mathrm{d}^2\boldsymbol{S} = -\frac{1}{2} \int_S \frac{\boldsymbol{B}}{B^3} \cdot \mathrm{d}^2\boldsymbol{S}$$

The integral $\int_S \left(\boldsymbol{B}/B^3\right) \cdot \mathrm{d}^2\boldsymbol{S}$ represents the *solid angle* that surface S subtends with respect to point $\boldsymbol{B} = 0$. Calling this angle Ω we get this amazingly beautiful result:

$$\gamma_\uparrow(\partial S) = -\frac{1}{2}\Omega$$

The phase shift experienced by a qubit in an eigenstate that is moved adiabatically along a closed trajectory in the $\boldsymbol{B}$ space is equal to $\pm 1/2$ (the sign depends on whether it is $|\uparrow\rangle$ or $|\downarrow\rangle$) times the solid angle subtended by the surface enclosed by the trajectory with respect to the $\boldsymbol{B} = 0$ point. *It is very seldom that we arrive at so startlingly elegant a result in physics or in mathematics.*

But what about our specific problem of a qubit in a rotating magnetic field. What is the actual *number*?

The number is easy to evaluate. Recall that our surface in this case is the hemisphere centered on $\boldsymbol{B} = 0$ of a constant radius B. Vector $\boldsymbol{B}$ in this case is parallel to the normal vector at each point of the surface, i.e., $\boldsymbol{B} = B\boldsymbol{n}$ and so is parallel to $\mathrm{d}^2\boldsymbol{S} = \boldsymbol{n}\mathrm{d}^2 S$ too. The integral becomes:

$$\gamma_\uparrow(\partial S) = -\frac{1}{2} \frac{1}{B^2} \int_S \boldsymbol{n} \cdot \boldsymbol{n}\mathrm{d}^2 S = -\frac{1}{2B^2} 2\pi B^2 = -\pi$$

and the phase factor γ is

$$e^{i\gamma_\uparrow} = e^{-i\pi} = -1$$

The geometric phase shift experienced by a qubit in adiabatically rotating magnetic field is equal to half of the solid angle subtended by the field trajectory in the $\boldsymbol{B}$ space.

and
$$e^{i\gamma_\downarrow} = e^{i\pi} = -1$$

For an arbitrary spinor $\mid \Psi \rangle = a \mid \uparrow \rangle + b \mid \downarrow \rangle$ the final result of this excursion is

$$\mid \Psi \rangle \rightarrow ae^{-iE_\uparrow \Delta t/\hbar}(-1) \mid \uparrow \rangle + be^{-iE_\downarrow \Delta t/\hbar}(-1) \mid \downarrow \rangle$$

where Δt is the time it takes to rotate the field. The time integrals in the dynamic phase part of the expression reduce to just $\exp\left(-iE_{\uparrow,\downarrow}\Delta t/\hbar\right)$, because B is constant (though $\boldsymbol{B}$ rotates slowly). This corresponds to Larmor precession around each direction of $\boldsymbol{B}$ the qubit goes through. Observe that the dynamic factor depends on time it takes to complete the excursion. On the other hand the γ factor depends on the solid angle only. This is why it is called the *geometric* phase factor. It is also called the Berry phase factor, or just the Berry phase for short.

But in this case it is clear that the geometric phase factor has no physical effect, because it vanishes in all expressions such as $\boldsymbol{\rho} = \mid \Psi \rangle\langle \Psi \mid$ or $|\langle \uparrow, \downarrow \mid \Psi \rangle|^2$.

But consider an excursion in the $\boldsymbol{B}$ space around a surface that subtends a solid angle Ω that is less than 2π. In the physical and in the $\boldsymbol{B}$ space this means that we rotate $\boldsymbol{B}$ conically rather than in a plane. This is shown in figure 4.8.
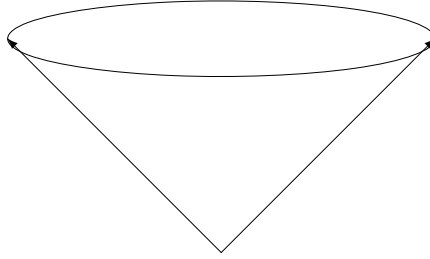
Figure 4.8: An excursion in the $\boldsymbol{B}$ space that results in a smaller solid angle $\Omega < 2\pi$.

In this case $e^{i\gamma} = e^{-i\Omega/2}$ and $e^{-i\gamma} = e^{i\Omega/2}$ no longer overlap and we get a physically observable effect.

Suppose we start our excursion with a qubit in the $\mid \rightarrow \rangle$ state

$$\mid \Psi(t = 0) \rangle = \frac{1}{\sqrt{2}} \left(\mid \uparrow \rangle + \mid \downarrow \rangle \right)$$

At the end of the excursion the qubit ends up in the state given by

$$\frac{1}{\sqrt{2}} \left(e^{i\mu B \Delta t/\hbar} e^{-i\Omega/2} \mid \uparrow \rangle + e^{-i\mu B \Delta t/\hbar} e^{i\Omega/2} \mid \downarrow \rangle \right)$$

What is going to be the effect of this change on the $\boldsymbol{r}$ vector that describes the qubit? The z component of $\boldsymbol{r}$ is not going to change, because

$$r^z = aa^* - bb^* = 0$$

But both r^x and r^y acquire an additional rotation due to $\boldsymbol{\gamma}$:

$$
\begin{aligned}
r^x &= ab^* + ba^* \\
&= \frac{1}{2}\left(e^{i(\mu B \Delta t/\hbar - \Omega/2)}e^{i(\mu B \Delta t/\hbar - \Omega/2)} + e^{-i(\mu B \Delta t/\hbar - \Omega/2)}e^{-i(\mu B \Delta t/\hbar - \Omega/2)}\right) \\
&= \frac{1}{2}\left(e^{i(2\mu B \Delta t/\hbar - \Omega)} + e^{-i(2\mu B \Delta t/\hbar - \Omega)}\right) \\
&= \cos(\omega_L \Delta t - \Omega)
\end{aligned}
$$

A qubit rotated by $\boldsymbol{B}(t)$ such that $\Omega < 2\pi$ acquires an angular lag of Ω behind the accumulated angle of the Larmor precesion.

where $\omega_L = 2\mu B/\hbar$ is the Larmor frequency, and

$$
\begin{aligned}
r^y &= i(ab^* - ba^*) \\
&= -\sin(\omega_L \Delta t - \Omega)
\end{aligned}
$$

We find that the qubit has accumulated an angular lag of Ω in the real physical space. The lag cancels itself away when $\Omega = 2\pi$.

4.10.4   Observing Berry phase experimentally

In the last 20 years or so, numerous papers have been published on the subject of Berry phase. A search on the automated e-print archives retrieves some 340 papers on Berry phase or on geometric phase. But very, very few of these are experimental papers that focus specifically on the demonstration of Berry phase in quantum systems.

Perhaps amongst the most elegant such demonstrations is the one by Richardson, Kilvington, Green and Lamoreaux [48] that was published in Physical Review Letters in 1988. Richardson and his colleagues from Institut Laue-Langevin in France and from the Rutherford-Appleton Laboratory in the UK used ultracold neutrons to demonstrate the Berry shift.

This method is very clean and the results obtained by the authors agree with theoretical predictions exceptionally well. This is very important because there are certain assumptions in the derivation of the Berry phase that may be questioned. The first one is the very notion of an adiabatic evolution of a qubit. The concept itself is not very clear and some authors proposed revisions. Alas, the surprising effect of such revisions was that the Berry phase disappeared, so that the revisions

themselves had to be revised in order to bring it back [41]. This development points to the crucial importance of experiment in physics. However convincing and mathematically elegant a derivation, there may be always something subtle in it that's missed or the underlying assumptions may not be correct. Experimental verification of a prediction provides us with certainty as to the proposed effect.

The experiment carried out by Richardson, Kilvington, Green and Lamoraux is additionally of great value and interest to us, because it demonstrates what is involved in precise manipulation and control of quantum systems to the level that would make quantum information processing feasible.

Storing neutrons in a chamber

The ultracold neutrons used in the experiment are neutrons that have been slowed down to less than 5 m/s. The neutrons are stored in a vacuum chamber lined with beryllium and beryllium oxide walls, since ultracold neutrons reflect from such walls with negligible loss for all incident angles. Nevertheless surface contamination and leaks in the valve that's used to fill and empty the chamber reduce the lifetime of neutrons dramatically, from nearly 10 minutes to only about 80 seconds. But 80 seconds is time a plenty in the quantum domain. We have seen in the quantronium example that the quantum state there decayed within about a *micro*second.

Generating and cooling neutrons

The neutrons are generated by irradiating a 25 litre container filled with liquid deuterium held at 25 K with fast neutrons emitted by a high-flux nuclear reactor. This irradiation procedure results in generation of ultra cold and *very* cold neutrons, which are then transported 10 m up through a nickel coated evacuated pipe. Nickel coating for the pipes is chosen again to limit neutron loss. The neutrons lose some of their energy in the process (to gravity) and are then pushed through a turbine that further slows the very cold neutrons thus doubling the number of ultra cold neutrons. At the output the utra cold neutron density is about $90/\text{cm}^3$.

Polarizing neutrons

The ultra cold neutrons are then passed through an 800 nm thick magnetically saturated polarizing foil made of cobalt-nickel alloy. Afterwards they are transported towards their 5 litre beryllium and beryllium oxide lined chamber along a copper-nickel alloy coated silica guide and pushed into the chamber through holes in the five-layer Mumetal magnetic shield. Mumetal is an 80% nickel-iron alloy that is specially designed for magnetic shielding applications.

Thermalizing and rotating neutrons in the chamber

It takes about 10 seconds to fill the chamber to the density of 10 neutrons per cubic centimeter, whereupon the neutron valve is closed. All neutrons in the chamber are initially polarized in the z direction by a 5 milligauss magnetic field, $\boldsymbol{B} = B_z^0 \boldsymbol{e}_z$. They are allowed to thermalize for 2 seconds and are then subjected to a temporally varying magnetic field for about 7.4 seconds, which is followed by another 2 seconds wait.

How is the field rotated

The varying magnetic field applied during the middle 7.4 seconds period is given

by

$$\boldsymbol{B} = aB_0\boldsymbol{e}_x \pm (1 + \epsilon)B_0 \cos\frac{2\pi t}{T} + B_0 \sin\frac{2\pi t}{T}$$

where a, ϵ, B_0 and T are constants and T is the time for one rotation.

The field rotates in the $\boldsymbol{e}_y \times \boldsymbol{e}_z$ plane and has a non-vanishing x component that lets the experimenters controll the solid angle Ω. Additionally ϵ is the ellipticity parameter that allows for variation of B over the path. The solid angle Ω that corresponds to the trajectory is given by

$$\Omega = 2\pi \left(1 \pm \frac{a}{\sqrt{1 + a^2}} \right)$$

where the sign $\pm$ depends on the sign of the magnetic moment and direction of the rotation.

After the excursion the neutron valve is opened and the same foil that was used *Extracting the* to polarize them in the first place is now used to filter them on the way out, thus *neutrons from* performing the measurement. An adiabatic spin flipper is used to filter both $|\uparrow\rangle$ *the chamber and* and $|\downarrow\rangle$ states, which are then counted in a helium proportional counter neutron *measuring them* detector for 10 seconds per state.

The magnetic field configuration used in the experiment is rotated by 90° com- *Why is the* pared to our examples discussed in the previous section. But this is done for a *rotation done in* good purpose. Recall that what we normally measure is the z component of the *the $\boldsymbol{e}_y \times \boldsymbol{e}_z$* spin, what we called p^0 and p^1 in chapter 2. If we were to perform the rotation in *plane* the $\boldsymbol{e}_x \times \boldsymbol{e}_y$ plane we would have to have a separate differently oriented device for measuring spins in the x and y directions. Furthermore we would have problems with preparation of the initial conditions too. As things are done in the experiment, when the neutrons enter the chamber they are polarized in the z direction and after thermalization their $p^0(t = 0)$ is 1 and their $r^z(t = 0) = 1$ too. They are then rotated in the $\boldsymbol{e}_y \times \boldsymbol{e}_z$ plane and their final angle of rotation, θ, can be read from $p^0 = (1 + r^z)/2$ and $p^1 = (1 - r^z)/2$, namely $r^z(t = T) = \cos\theta = p^0(T) - p^1(T)$, as measured by the same polarizing foil when the neutrons are extracted from the chamber.

The final angle of rotation contains the accumulated dynamic phase and the geometric phase, i.e.,

$$\theta_\uparrow = 2\mu \int_0^T B(t)\mathrm{d}t - \Omega \tag{4.56}$$

and $\theta_\downarrow = -\theta_\uparrow$.

For a multiple number of revolutions of the field, N, both the dynamic and the geometric phases accumulate so in this case we get $\theta_\uparrow = N\left(2\mu \int_0^T B(t)\mathrm{d}t - \Omega \right)$.

How was the magnetic field generated and controlled

The magnetic field was generated by running current through three sets of coils, which were placed within the magnetic shield, but outside the vacuum chamber containing the neutrons. The coils were perpendicular to each other with accuracy of better than $2°$ and were calibrated to within 0.1% accuracy. An analog computer was used to control the currents. In turn a timer and a zero-crossing switch were used to control the computer, so that exactly one full rotation, or a multiple thereof, could be generated.

Sources of depolarization

In spite of all the precautions and the use of special materials there was a residual magnetic field of about $10\,\mu G$ and a non-vanishing gradient in the neutron chamber. This residual field was strong enough to depolarise the neutrons for small values of B_0. The presence of the residual field limited the duration of the rotation T and set a lower limit to B_0.

Another problem was caused by an aluminum can in which the neutron chamber was enclosed. Varying $\boldsymbol{B}$ too quickly would generate eddy currents in the can, which would rapidly depolarize the neutrons too. So T could not have been made too short. The choice of $T \approx 7.4\,\mathrm{s}$ resulted from this restriction.

Collecting neutron counts

The measurements proceeded as follows. A given set of parameters N, a and ϵ would be fixed. For this set of parameters the spin-up and spin-down counts were collected as a function of B_0. Each chamber fill and store cycle would yield a certain number of counts. This was being repeated until between 60,000 and 70,000 counts were collected for *each* (N, a, ϵ, B_0) tuple.

Calibrating the instrument and weighing the counts

The collected counts had to be weighed to correct for the fact that spin-up neutrons were counted first, and while they were being counted, the spin-down neutrons were stored in the guide and reflected off the polarizer, thus suffering additional depolarization. Other sources of depolarization had to be included in final data analysis too. This is normally done by *calibrating* the system against known neutron configurations obtained, e.g., by filling the chamber with polarized neutrons, thermalizing them, doing nothing to them for 7.4 s, then emptying the chamber and counting the neutrons.

Experimental results - the Berry phase is clearly seen and accurately measured

In the end, after many days of collecting and processing data[6], the experimenters arrived at the numbers shown here in table 4.10.4.

The agreement between Berry's predictions and observed values of the additional angle due to the geometric phase shift is very good. Ellipticity of the orbit has no effect as long as the solid angle in the $\boldsymbol{B}$ space remains unchanged. The Berry angle is clearly accumulative – this can be seen by comparing data for various values of N.

---

[6]What does this tell us about the feasibility of using quantum systems for computations?

			calculated	observed	
a	ϵ	N	$\Omega/2\pi$	$\Omega_u/2\pi$	$\Omega_d/2\pi$
0.000	0.00	1	1.000	1.00 ± 0.01	1.00 ± 0.02
0.000	0.25	1	1.000	1.00 ± 0.03	0.99 ± 0.05
0.000	0.62	1	1.000	1.01 ± 0.03	1.00 ± 0.05
0.000	0.00	2	2.000	2.00 ± 0.03	1.97 ± 0.06
0.000	0.00	3	3.000	2.87 ± 0.15	2.89 ± 0.15
0.268	0.00	1	1.259	1.28 ± 0.01	1.26 ± 0.03
0.577	0.00	1	1.500	1.52 ± 0.02	1.51 ± 0.03
0.577	0.00	2	3.000	3.00 ± 0.03	2.99 ± 0.05
1.000	0.00	1	1.707	1.68 ± 0.01	1.69 ± 0.02
1.732	0.00	1	1.866	1.74 ± 0.15	1.72 ± 0.40
3.732	0.00	1	1.966	1.97 ± 0.01	1.98 ± 0.02

Table 4.1: Results of the experiment. a defines the solid angle Ω, ϵ is the ellipticity, N is the number of rotations, Ω is the theoretical value of the Berry angle, see equation (4.56), which should be equal to the solid angle and to the solid angle alone, Ω_u is the measured value of the Berry angle obtained by counting spin-up neutrons and Ω_d is the measured value of the Berry angle obtained by counting spin-down neutrons. Observe increased depolarization for $N = 3$ and the lack of effect for $\epsilon \neq 0$.

4.10.5 Berry phase gates

Berry phase has been proposed as an additional mechanism for processing quantum information. The device discussed in the previous section constitutes an example of a Berry phase gate. Jones, Vedral, Ekert and Castagnoli even demonstrated a *conditional* Berry phase gate using nuclear magnetic resonance [33]. Yau, De Poortere and Shayegan detected signs of Berry phase in oscillations of the resistance of a mesoscopic gallium arsenide ring embedded in a magnetic field – although this could have been demonstrated only after the measured spectra were compared with simulation results [54]. This gives us hope that we may see a more direct evidence for Berry phase in quantum electronic devices one day. *(Examples of Berry phase gates)*

But this idea is not without its problems.

The first difficulty is that the Berry phase is always mixed with Larmor precession in a way that Rabi oscillations aren't. Recall that Rabi oscillations affect p^0/p^1, whereas Larmor precession leaves p^0 and p^1 intact, unless the background field is rotated by $90°$ as has been done in the Richardson's experiment. This makes it possible to design a computational system based on Rabi oscillations that utilizes *(Berry phase is mixed with Larmor precesion)*

p^0 and p^1 but ignores p^2 and p^3. Indeed most quantum algorithms developed so far and their experimental demonstrations do just this.

Berry phase gates are imprecise

The second problem is lack of precision. It is very difficult to rotate the whole external magnetic field system by precisely one revolution and under exactly the solid angle that we need. The Richardson's experiment is extraordinarily precise for quantum mechanics, but even here we have seen as much as 23% errors for some (N, a, ϵ, B_0) tuples and for statistical ensembles of some 70,000 events per tuple. Berry shift that is seen in the Yau's gallium arsenide ring is so far out that the experimenters could only claim *qualitative* agreement with Berry's predictions. The idea of actually *moving* qubits physically in 3-D space, rather than subjecting them to a passive excursion as it has been done in the Richardson's experiment is probably completely unrealistic, since touching and pushing a qubit, by whatever means, would most likely destroy its quantum state to begin with.

Individual qubits cannot be addressed with Berry gates

Then we have the problem of addressing individual qubits in a quantum register. The register may be a molecule, or a collection of atoms trapped in an optical lattice or something else. Rotating magnetic field around the register would perform the operation on all its qubits, not just on a selected qubit. The usefulness of such an operation is likely to be very limited. On the other hand the Rabi oscillations mechanism lets us talk to individual qubits on private channels. If a register is a specially constructed molecule, atoms in various locations within the molecule are sensitive to different Rabi frequencies because of the so called chemical shifts. By sending signals on these frequencies we can address individual qubits and not have other qubits eavesdrop on the communication. Similarly, we can *read* individual qubits by tuning the receivers to chemically shifted Rabi frequencies.

This last problem could be overcome in quantum electronic circuits if every qubit could be equipped in its own local "magnetic field" circuitry, the way it has been done, for example, in the quantronium.

Berry gates are slow

Last but not least we have the issue of adiabatic transport. Berry phase equation does not work for non-adiabatic transport. If a qubit manipulation is too fast we have to solve the Schrödinger equation exactly without making the adiabatic assumption that an eigenvector remains an eigenvector throughout the whole excursion. For a transport to be adiabatic it must be slow. But a slow qubit manipulation means a slow gate and inevitable high depolarization rate while the gate is being traversed.

In summary, rather than thinking of Berry phase as a possible gate mechanism, we may have to think about it as a yet another parasitic effect, alongside with Larmor precesion, that has to be kept under control while quantum information is being processed.

5 The Biqubit

5.1 Entangled states

Whereas a single, isolated qubit is mathematically equivalent to a classical magnetic dipole—be it with some read-out and statistical complications—a system of two qubits, a *biqubit*, is equivalent to two classical magnetic dipoles occasionally only. At other times it displays a rather puzzling behaviour that had stirred a great deal of theoretical and experimental investigations towards the end of the XXth century and from which the ideas of quantum computing eventually arose.

But let us begin by considering two *separate* qubits first, i.e., qubits that *are* equivalent to two classical magnetic dipoles.

Using the elementary laws of probability calculus, as we have discussed in section 1.9, page 37, we would describe the qubits in terms of a tensor product of their fiducial vectors, i.e., we would say that the state of the system comprising qubits A and B is

$$\boldsymbol{p}_A \otimes \boldsymbol{p}_B$$

Measurements on this system would then be expressed by a tensor product of two forms: the first one, $\boldsymbol{\omega}_A$, describing a measurement on $\boldsymbol{p}_A$ and the second one, $\boldsymbol{\omega}_B$, describing a measurement on $\boldsymbol{p}_B$. And so, for example, the probability that qubit A is in state $|\uparrow\rangle$ is

$$p^0 = \langle \boldsymbol{\omega}_A^0, \boldsymbol{p}_A \rangle,$$

where $\boldsymbol{\omega}_A^0$ is a canonical form in the space of qubit A.

> To remind you, "canonical" here means that it simply extracts the 0th component of vector $\boldsymbol{p}_A$. See equations 2.15 on page 64 and 2.16 on page 64.

The probability that qubit B is in state $|\rightarrow\rangle$ is

$$p^2 = \langle \boldsymbol{\omega}_B^2, \boldsymbol{p}_B \rangle,$$

and the probability that qubit A is in state $|\uparrow\rangle$ while qubit B is in state $|\rightarrow\rangle$ – of all other possible two-qubit combinations – is

$$p^0 p^2 = \langle \boldsymbol{\omega}_A^0, \boldsymbol{p}_A \rangle \langle \boldsymbol{\omega}_B^2, \boldsymbol{p}_B \rangle = \langle \boldsymbol{\omega}_A^0 \otimes \boldsymbol{\omega}_B^2, \boldsymbol{p}_A \otimes \boldsymbol{p}_B \rangle$$

The two-qubit energy form is somewhat more complicated, because energy is an additive quantity, i.e., energy of a two-qubit system is a sum of energies of the two qubits, as long as they don't interact with each other.

This can be captured in the following way. Recall that

$$\langle \varsigma^i, \varsigma_j \rangle = 2\delta^i{}_j, \quad i,j = 1, x, y, z$$

therefore

$$\langle \varsigma^1, \boldsymbol{p} \rangle = \langle \varsigma^1, \frac{1}{2} \left(\varsigma_1 + r^x \varsigma_x + r^y \varsigma_y + r^z \varsigma_z \right) \rangle = 1$$

Consequently if $\boldsymbol{\eta}_A$ and $\boldsymbol{\eta}_B$ are the energy forms acting on the fiducial vector of qubits A and B respectively, the energy form for a system of two separate non-interacting qubits is:

$$\boldsymbol{\eta}_A \otimes \varsigma_B^1 + \varsigma_A^1 \otimes \boldsymbol{\eta}_B$$

What if the qubits do interact with each other? Then the energy form may have an additional term that couples to both qubits simultaneously

$$\boldsymbol{\eta}_A \otimes \varsigma_B^1 + \varsigma_A^1 \otimes \boldsymbol{\eta}_B + \boldsymbol{\eta}_{AB}$$

where $\boldsymbol{\eta}_{AB}$ is a form of rank two, i.e., a form that acts on objects such as $\boldsymbol{p}_A \otimes \boldsymbol{p}_B$.

When evaluating biqubit expressions, we must always remember that only forms operating on qubit A can be applied to this qubit. Expressions such as

$$\langle \boldsymbol{\eta}_A, \boldsymbol{p}_B \rangle$$

make no sense, because form $\boldsymbol{\eta}_A$ does *not* operate in the space of states of qubit B.

We can switch between fiducial and quaternion formalisms by converting qubit probability vectors $\boldsymbol{p}$ to the corresponding quaternions $\boldsymbol{\rho}$, keeping at the same time the tensor product symbol in place:

$$\boldsymbol{p}_A \otimes \boldsymbol{p}_B \rightarrow \boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B$$

And then we can substitute Pauli matrices in place of quaternion units $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$, and $\boldsymbol{\sigma}_z$, so that the tensor product of quaternions becomes the tensor product of Pauli matrices.

We *must not* yield to the temptation of just multiplying $\boldsymbol{\rho}_A$ by $\boldsymbol{\rho}_B$ – regardless of whether the sigmas are thought of as quaternions or Pauli matrices. The *product in waiting* must wait and the two separate vector spaces are mapped on two separate quaternion spaces and these in turn are mapped on two separate spaces of 2×2 complex matrices.

To emphasize this we're going to add subscripts A and B to the sigmas as well

$$\boldsymbol{p}_A \quad \rightarrow \quad \frac{1}{2} \left(\mathbf{1}_A + r_A^x \boldsymbol{\sigma}_{xA} + r_A^y \boldsymbol{\sigma}_{yA} + r_A^z \boldsymbol{\sigma}_{zA} \right)$$

$$\boldsymbol{p}_B \quad \rightarrow \quad \frac{1}{2} \left(\mathbf{1}_B + r_B^x \boldsymbol{\sigma}_{xB} + r_B^y \boldsymbol{\sigma}_{yB} + r_B^z \boldsymbol{\sigma}_{zB} \right)$$

Matrices $\boldsymbol{\sigma}_{iA}$ and $\boldsymbol{\sigma}_{iB}$ look the same as normal Pauli matrices $\boldsymbol{\sigma}_i$, but operate in different spaces. This must be emphasized *ad nauseam*.

Having unpacked the sigmas onto 2×2 matrices, we can go further and express our qubit states in terms of Hilbert space vectors, $\mid \Phi \rangle_A$ and $\mid \Psi \rangle_B$ such that:

$$
\begin{aligned}
\mid \Phi \rangle_A \otimes {}_A\langle \Phi \mid &= \boldsymbol{\rho}_A, \quad \text{and} \\
\mid \Psi \rangle_B \otimes {}_B\langle \Psi \mid &= \boldsymbol{\rho}_B
\end{aligned}
$$

This is going to work *only* when both constitutent states $\boldsymbol{p}_A$ and $\boldsymbol{p}_B$ are *pure*. For such pure states we have

$$
\begin{aligned}
\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B &= (\mid \Phi \rangle_A \otimes {}_A\langle \Phi \mid) \otimes (\mid \Psi \rangle_B \otimes {}_B\langle \Psi \mid) \\
&= (\mid \Phi \rangle_A \otimes \mid \Psi \rangle_B) \otimes ({}_A\langle \Phi \mid \otimes {}_B\langle \Psi \mid)
\end{aligned}
$$

So we end up with tensor products of two Hilbert space vectors or forms representing a biqubit system.

In summary, our chain of mappings from measurable probabilities for a system of two qubits in pure states to highly abstract (though convenient) states in the Hilbert space looks as follows:

$$
\boldsymbol{p}_A \otimes \boldsymbol{p}_B \rightarrow \boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B \rightarrow \mid \Phi \rangle_A \otimes \mid \Psi \rangle_B
$$

Let

$$
\begin{aligned}
\mid \Phi \rangle_A &= a \mid\uparrow\rangle_A + b \mid\downarrow\rangle_A, \quad \text{and} \\
\mid \Psi \rangle_B &= c \mid\uparrow\rangle_B + d \mid\downarrow\rangle_B
\end{aligned}
$$

where a, b, c and d are four complex numbers such that $aa^* + bb^* = 1$ and $cc^* + dd^* = 1$. Then

$$
\begin{aligned}
\mid \Phi \rangle_A \otimes \mid \Psi \rangle_B = {}&ac \mid\uparrow\rangle_A \otimes \mid\uparrow\rangle_B + ad \mid\uparrow\rangle_A \otimes \mid\downarrow\rangle_B \\
&+ bc \mid\downarrow\rangle_A \otimes \mid\uparrow\rangle_B + bd \mid\downarrow\rangle_A \otimes \mid\downarrow\rangle_B
\end{aligned}
$$

Observe that normalization conditions imposed on the Hilbert space states of individual qubits result in the following normalization condition of the biqubit state:

$$
\begin{aligned}
ac(ac)^* &+ ad(ad)^* + bc(bc)^* + bd(bd)^* \\
&= aa^*(cc^* + dd^*) + bb^*(cc^* + dd^*) \\
&= aa^* + bb^* = 1
\end{aligned}
$$

In summary, we find that a biqubit state

$$\alpha \left|\uparrow\right\rangle_A \otimes \left|\uparrow\right\rangle_B + \beta \left|\uparrow\right\rangle_A \otimes \left|\downarrow\right\rangle_B + \gamma \left|\downarrow\right\rangle_A \otimes \left|\uparrow\right\rangle_B + \delta \left|\downarrow\right\rangle_A \otimes \left|\downarrow\right\rangle_B$$

such that

$$\begin{aligned}
\alpha &= ac \\
\beta &= ad \\
\gamma &= bc \\
\delta &= bd
\end{aligned}$$

where $aa^* + bb^* = cc^* + dd^* = 1$ represents two *separate* qubits.

Dividing the first equation by the second one and then the third equation by the fourth one yields:

$$\alpha/\beta = c/d = \gamma/\delta$$

In turn, dividing the first equation by the third one and then the second equation by the fourth one yields:

$$\alpha/\gamma = a/b = \beta/\delta$$

Both equations are in fact the same and equivalent to

$$\alpha\delta - \beta\gamma = 0$$

or

$$\det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = 0 \tag{5.1}$$

This is a simple criterion that we can use to check if a given *pure* (because here we're within the unitary formalism) biqubit state can be separated into two independent qubits at all. To be more precise this is a *necessary* though not *sufficient* condition. But *necessary* is good enough if we want to prove that a given pure biqubit state is not separable.

To see how this criterion works consider the following biqubit state:

$$\left|\Psi^-\right\rangle_{AB} = \frac{1}{\sqrt{2}} \left(\left|\uparrow\right\rangle_A \otimes \left|\downarrow\right\rangle_B - \left|\downarrow\right\rangle_A \otimes \left|\uparrow\right\rangle_B \right) \tag{5.2}$$

For this state $\alpha = \delta = 0$ but $\beta = 1/\sqrt{2} = -\gamma$. The state is normalized because $\beta^2 + \gamma^2 = 1$ but

$$\det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = -\beta\gamma = \frac{1}{2}$$

So this is an example of a pure biqubit state that *cannot* be split into two separate qubits.

Can such states exist? If so, what do they mean and what is their fiducial, i.e., observable representation?

In order to arrive at the fiducial representation of this biqubit state we're going to convert it to a density operator and then to a quaternion representation.

$$
\begin{aligned}
\boldsymbol{\rho}_{AB} &= |\Psi^-\rangle_{AB} \otimes {}_{AB}\langle\Psi^-| \\
&= \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_A \otimes |\downarrow\rangle_B - |\downarrow\rangle_A \otimes |\uparrow\rangle_B \right) \otimes \frac{1}{\sqrt{2}} \left({}_A\langle\uparrow| \otimes {}_B\langle\downarrow| - {}_A\langle\downarrow| \otimes {}_B\langle\uparrow| \right) \\
&= \frac{1}{2} \Big(\left(|\uparrow\rangle_A \otimes {}_A\langle\uparrow|\right) \otimes \left(|\downarrow\rangle_B \otimes {}_B\langle\downarrow|\right) - \left(|\uparrow\rangle_A \otimes {}_A\langle\downarrow|\right) \otimes \left(|\downarrow\rangle_B \otimes {}_B\langle\uparrow|\right) \\
&\quad - \left(|\downarrow\rangle_A \otimes {}_A\langle\uparrow|\right) \otimes \left(|\uparrow\rangle_B \otimes {}_B\langle\downarrow|\right) + \left(|\downarrow\rangle_A \otimes {}_A\langle\downarrow|\right) \otimes \left(|\uparrow\rangle_B \otimes {}_B\langle\uparrow|\right) \Big)
\end{aligned}
$$

Invoking expressions we have arrived at in section 4.2, page 114, we can convert this readily to 2×2 matrices:

$$
\begin{aligned}
\boldsymbol{\rho}_{AB} = \frac{1}{2} \Bigg(&\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_A \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_B - \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_A \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_B \\
&- \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_A \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_B + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_A \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_B \Bigg)
\end{aligned}
$$

Now we use equations 4.8 on page 112 to express the matrices in terms of Pauli matrices:

$$
\begin{aligned}
\boldsymbol{\rho}_{AB} = \frac{1}{2} \Bigg(&\frac{1}{2}\left(\mathbf{1}_A + \boldsymbol{\sigma}_{zA}\right) \otimes \frac{1}{2}\left(\mathbf{1}_B - \boldsymbol{\sigma}_{zB}\right) \\
&- \frac{1}{2}\left(\boldsymbol{\sigma}_{xA} + i\boldsymbol{\sigma}_{yA}\right) \otimes \frac{1}{2}\left(\boldsymbol{\sigma}_{xB} - i\boldsymbol{\sigma}_{yB}\right) \\
&- \frac{1}{2}\left(\boldsymbol{\sigma}_{xA} - i\boldsymbol{\sigma}_{yA}\right) \otimes \frac{1}{2}\left(\boldsymbol{\sigma}_{xB} + i\boldsymbol{\sigma}_{yB}\right) \\
&+ \frac{1}{2}\left(\mathbf{1}_A - \boldsymbol{\sigma}_{zA}\right) \otimes \frac{1}{2}\left(\mathbf{1}_B + \boldsymbol{\sigma}_{zB}\right) \Bigg)
\end{aligned}
$$

which simplifies eventually to

$$
\boldsymbol{\rho}_{AB} = \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B - \boldsymbol{\sigma}_{xA} \otimes \boldsymbol{\sigma}_{xB} - \boldsymbol{\sigma}_{yA} \otimes \boldsymbol{\sigma}_{yB} - \boldsymbol{\sigma}_{zA} \otimes \boldsymbol{\sigma}_{zB}\right)
$$

At this stage we can convert this to measurable probabilities by replacing Pauli sigmas with Pauli varsigmas:

$$\boldsymbol{p}_{AB} = \frac{1}{4} \left(\boldsymbol{\varsigma}_{1A} \otimes \boldsymbol{\varsigma}_{1B} - \boldsymbol{\varsigma}_{xA} \otimes \boldsymbol{\varsigma}_{xB} - \boldsymbol{\varsigma}_{yA} \otimes \boldsymbol{\varsigma}_{yB} - \boldsymbol{\varsigma}_{zA} \otimes \boldsymbol{\varsigma}_{zB} \right) \qquad (5.3)$$

And this is our matrix of measurable probabilities:

$$\boldsymbol{p}_{AB} = \begin{pmatrix} p^{00} & p^{01} & p^{02} & p^{03} \\ p^{10} & p^{11} & p^{12} & p^{13} \\ p^{20} & p^{21} & p^{22} & p^{23} \\ p^{30} & p^{31} & p^{32} & p^{33} \end{pmatrix} = \begin{pmatrix} p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\ p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\ p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\ p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes} \end{pmatrix}$$

It has 16 entries and it specifies the biqubit system *entirely* in terms of probabilities of detecting qubit A's "spin" against one "direction" and qubit B's "spin" against some other "direction" at the same time - with "directions" for both qubits being $\pm\boldsymbol{e}_z$ ($\uparrow$, $\downarrow$), $\boldsymbol{e}_x$ ($\rightarrow$) and $\boldsymbol{e}_y$ ($\otimes$). The measurements on each qubit are carried out the same way as before, but this time it is *not* enough to measure each qubit separately. To fully characterize the state, we need 16 probabilities of two specific events happening simultaneously.

How simultaneous do they have to be? Simultaneous enough so that, say, qubit B does not have a chance of interacting with the environment after qubit A has been measured.

How to get the actual numbers, p^{ij}, from the varsigma expression? This is easier than it may seem at first glance.

Recall that

$$p^{ij} = \langle \boldsymbol{\omega}^i \otimes \boldsymbol{\omega}^j, \boldsymbol{p}_{AB} \rangle$$

where $\boldsymbol{\omega}^i$ are the canonical forms. The canonical forms are not dual to Pauli vectors, but they are dual to canonical vectors $\boldsymbol{e}_i$ and we know how to express Pauli vectors in terms of canonical vectors, because equations 2.11 on page 62 specify the procedure, namely:

$$\begin{aligned} \boldsymbol{\varsigma}_1 &= \boldsymbol{e}_0 + \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 \\ \boldsymbol{\varsigma}_x &= \boldsymbol{e}_2 \\ \boldsymbol{\varsigma}_y &= \boldsymbol{e}_3 \\ \boldsymbol{\varsigma}_z &= \boldsymbol{e}_0 - \boldsymbol{e}_1 \end{aligned}$$

So the trick here is to replace the varsigmas in equation 5.3 with canonical vectors, and this will flush out the actual probabilities right away.

Having made the corresponding substitutions we obtain

$$
\begin{aligned}
\boldsymbol{p}_{AB} \;=\; \frac{1}{4}\Big(& (\boldsymbol{e}_{0A}+\boldsymbol{e}_{1A}+\boldsymbol{e}_{2A}+\boldsymbol{e}_{3A})\otimes(\boldsymbol{e}_{0B}+\boldsymbol{e}_{1B}+\boldsymbol{e}_{2B}+\boldsymbol{e}_{3B}) \\
& -\boldsymbol{e}_{2A}\otimes\boldsymbol{e}_{2B}-\boldsymbol{e}_{3A}\otimes\boldsymbol{e}_{3B} \\
& -(\boldsymbol{e}_{0A}-\boldsymbol{e}_{1A})\otimes(\boldsymbol{e}_{0B}-\boldsymbol{e}_{1B})\Big)
\end{aligned}
$$

The first term corresponds to a 4×4 matrix of ones:

$$
(\boldsymbol{e}_{0A}+\boldsymbol{e}_{1A}+\boldsymbol{e}_{2A}+\boldsymbol{e}_{3A})\otimes(\boldsymbol{e}_{0B}+\boldsymbol{e}_{1B}+\boldsymbol{e}_{2B}+\boldsymbol{e}_{3B})
$$
$$
=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}
$$

The second and third terms correspond to matrices of zeros with 1 in the $(2,2)$ and $(3,3)$ positions respectively:

$$
\boldsymbol{e}_{2A}\otimes\boldsymbol{e}_{2B}=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\quad \boldsymbol{e}_{3A}\otimes\boldsymbol{e}_{3B}=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

and the last term corresponds to a matrix that looks as follows:

$$
(\boldsymbol{e}_{0A}-\boldsymbol{e}_{1A})\otimes(\boldsymbol{e}_{0B}-\boldsymbol{e}_{1B})
$$
$$
=\boldsymbol{e}_{0A}\otimes\boldsymbol{e}_{0B}-\boldsymbol{e}_{0A}\otimes\boldsymbol{e}_{1B}-\boldsymbol{e}_{1A}\otimes\boldsymbol{e}_{0B}+\boldsymbol{e}_{1A}\otimes\boldsymbol{e}_{1B}
$$
$$
=\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

Combining the matrices yields

$$
\boldsymbol{p}_{AB} \;=\; \frac{1}{4}\left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}-\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right.
$$
$$
\left.-\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}-\begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\right)
$$

$$
= \quad \frac{1}{4} \begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\ p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\ p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\ p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes} \end{pmatrix}
$$

Now we can finally get down to physics. And the physics that emerges from this matrix is most peculiar. The first thing to notice is that

$$
p^{\uparrow\uparrow} = p^{\downarrow\downarrow} = p^{\rightarrow\rightarrow} = p^{\otimes\otimes} = 0
$$

If the biqubit is made of two neutrons and has been prepared in this special state, the probability of finding the neutrons aligned, i.e., both in the $|\uparrow\rangle$ state, or both in the $|\downarrow\rangle$ state, or both in the $|\rightarrow\rangle$ state or both in the $|\otimes\rangle$ state is... zero. Regardless of how we orient the measuring apparatuses, the qubits (neutrons in this case) always come out pointing in the opposite directions – even if the measuring polarizers are far away from each other, but this only as long as the neutrons are still described by $\left(|\uparrow\rangle_A \otimes |\downarrow\rangle_B - |\downarrow\rangle_A \otimes |\uparrow\rangle_B\right)/\sqrt{2}$, which is not going to be forever, because this peculiar biqubit state is going to depolarize faster even than single-qubit polarized states.

The upper left corner of matrix $\boldsymbol{p}_{AB}$ is normalized. States $|\uparrow\rangle_A \otimes |\uparrow\rangle_B$, $|\uparrow\rangle_A \otimes |\downarrow\rangle_B$, $|\downarrow\rangle_A \otimes |\uparrow\rangle_B$ and $|\downarrow\rangle_A \otimes |\downarrow\rangle_B$ constitute the physical basis of the system and

$$
p^{\uparrow\uparrow} + p^{\uparrow\downarrow} + p^{\downarrow\uparrow} + p^{\downarrow\downarrow} = 1
$$

The remaining entries in the matrix describe measurements against directions that are perpendicular to each other, for example, if qubit A is measured against $\boldsymbol{e}_x$ and qubit B is measured against $\boldsymbol{e}_y$, then the probability associated with the outcome is $p^{\rightarrow\otimes} = 1/4$. This result is consistent with the assumption that both qubits in this state must always be found pointing in the opposite directions. So, for example, if qubit A is found pointing in the $\boldsymbol{e}_x$ direction, which on the whole is going to be $1/2$ of all cases associated with this measurement, then qubit B should point in the $-\boldsymbol{e}_x$ direction, but when measured against the $\boldsymbol{e}_y$ direction, half of all qubits B will emerge pointing in the $\boldsymbol{e}_y$ direction and the other half pointing in the $-\boldsymbol{e}_y$ direction. And so the probabilty of finding that qubit A points in the $\boldsymbol{e}_x$ direction and qubit B points in the $\boldsymbol{e}_y$ direction is $1/2 \times 1/2 = 1/4$.

There are three other similar states, whose probability matrix can be computed the same way. They are

$$
| \Psi^+\rangle_{AB} \quad = \quad \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_A \otimes |\downarrow\rangle_B + |\downarrow\rangle_A \otimes |\uparrow\rangle_A\right) \tag{5.4}
$$

$$\boldsymbol{\rho}_{AB} \;=\; \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B + \boldsymbol{\sigma}_{xA}\otimes\boldsymbol{\sigma}_{xB} + \boldsymbol{\sigma}_{yA}\otimes\boldsymbol{\sigma}_{yB} - \boldsymbol{\sigma}_{zA}\otimes\boldsymbol{\sigma}_{zB}\right)$$

$$\boldsymbol{p}_{AB} \;=\; \begin{pmatrix} p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\ p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\ p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\ p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes} \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

This state describes a biqubit system of "total spin 1" but with a zero component in the $\boldsymbol{e}_z$ direction. When projected on either $\boldsymbol{e}_x$ or $\boldsymbol{e}_y$ both 1/2-spins of the biqubit align and add up. But when projected on $\boldsymbol{e}_z$ the spins counter-align, and so the projection of the total spin on this direction is zero.

Recall that $\left(|\uparrow\rangle_A \otimes |\downarrow\rangle_B - |\downarrow\rangle_A \otimes |\uparrow\rangle_B\right)/\sqrt{2}$ was characterized by 1/2-spins counter- ▮
aligning in any direction – that state was therefore a "total spin 0" state.

The next state is:

$$|\,\Phi^-\rangle_{AB} \;=\; \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_A \otimes |\uparrow\rangle_B - |\downarrow\rangle_A \otimes |\downarrow\rangle_A\right) \tag{5.5}$$

$$\boldsymbol{\rho}_{AB} \;=\; \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B - \boldsymbol{\sigma}_{xA}\otimes\boldsymbol{\sigma}_{xB} + \boldsymbol{\sigma}_{yA}\otimes\boldsymbol{\sigma}_{yB} + \boldsymbol{\sigma}_{zA}\otimes\boldsymbol{\sigma}_{zB}\right)$$

$$\boldsymbol{p}_{AB} \;=\; \begin{pmatrix} p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\ p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\ p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\ p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes} \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

This state is similar to $\left(|\uparrow\rangle_A \otimes |\downarrow\rangle_B + |\downarrow\rangle_A \otimes |\uparrow\rangle_B\right)/\sqrt{2}$, meaning that this is also a "total spin 1" state, but this time it is the $\boldsymbol{e}_x$ component of the spin that is missing. When projected on this direction, both spins counter-align, but they align when projected on $\boldsymbol{e}_z$ or $\boldsymbol{e}_x$.

And finally we have

$$|\,\Phi^+\rangle_{AB} \;=\; \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_A \otimes |\uparrow\rangle_B + |\downarrow\rangle_A \otimes |\downarrow\rangle_A\right) \tag{5.6}$$

$$\boldsymbol{\rho}_{AB} \;=\; \frac{1}{4}\left(\mathbf{1}_A \otimes \mathbf{1}_B + \boldsymbol{\sigma}_{xA}\otimes\boldsymbol{\sigma}_{xB} - \boldsymbol{\sigma}_{yA}\otimes\boldsymbol{\sigma}_{yB} + \boldsymbol{\sigma}_{zA}\otimes\boldsymbol{\sigma}_{zB}\right)$$

$$\boldsymbol{p}_{AB} \;=\; \begin{pmatrix} p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\ p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\ p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\ p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes} \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

which is like the other two states above, i.e., a "total spin 1" state, with the $\boldsymbol{e}_y$ component missing.

Observe how adept the unitary formalism is at hiding the full physical characterizations of these states. We have to go through a number of quite complicated transformations to arrive at the probability matrices.

The density matrix (or quaternion) formalism is somewhat better. With a little practice one can see the actual probability matrices hiding inside $\boldsymbol{\rho}_{AB}$.

The fiducial formalism gives us the probability matrices explicitly. Alas, for three qubits $\boldsymbol{p}_{ABC}$ is going to be a cube, and a hypercube for four qubits and at this stage the fiducial representation quickly becomes too complex.

States $\mid \Psi^- \rangle_{AB}$ (equation (5.2)), $\mid \Psi^+ \rangle_{AB}$ (equation (5.4)), $\mid \Phi^- \rangle_{AB}$ (equation (5.5)) and $\mid \Phi^+ \rangle_{AB}$ (equation (5.6)) are said to be *entangled*. The verb *to entangle* means (1) *to wrap or twist together* – there is indeed a degree of togetherness in these states that, as we shall see later, cannot be explained by naive classical physics reasoning based on the concept of *local realism*[1], and leads to amusing paradoxes. It also means (2) *to involve in a perplexing or troublesome situation* – and this meaning is right on the spot too. The paradoxes mentioned above have been perplexing physics community ever since John Stewart Bell (1928-1990) came up with their concise mathematical characterization in 1964[2] [3], [4]. To make things worse they are not just theoretical paradoxes to be contemplated by arm-chair philosophers. They have all been confirmed by elaborate experiments [18], and have become fundamental to quantum computing.

In memory of John Bell, states $\mid \Psi^- \rangle_{AB}$, $\mid \Psi^+ \rangle_{AB}$, $\mid \Phi^- \rangle_{AB}$ and $\mid \Phi^+ \rangle_{AB}$ are called *Bell states*.

How do Bell states differ from separable biqubit states? Well, we already know that they can't be separated from the unitary formalism, but how does this manifest on the fiducial formalism level?

A biqubit state that is made of two *separate* qubits, each in its own well defined state that may be a mixture, has the following fiducial representation:

$$
\begin{aligned}
\boldsymbol{p}_{AB} &= \boldsymbol{p}_A \otimes \boldsymbol{p}_B \\
&= \frac{1}{4} \left(\boldsymbol{\varsigma}_{1A} + \sum_{i=x,y,z} r_A^i \boldsymbol{\varsigma}_{iA} \right) \otimes \left(\boldsymbol{\varsigma}_{1B} + \sum_{j=x,y,z} r_B^j \boldsymbol{\varsigma}_{jB} \right)
\end{aligned}
$$

---

[1] . . . but it can be explained by less naive classical physics reasoning that abandons the locality assumption [4] [8] [14].

[2] The paradoxes themselves go further back to Einstein, Podolsky and Rosen, who discussed one such paradox in their paper in 1935 [16]. But EPR, as the trio is affectionately called, did not produce an experimentally verifiable formula that could be used to check which way the chips fall.

$$= \frac{1}{4} \left(\varsigma_{1A} \otimes \varsigma_{1B} + \sum_{j=x,y,z} r_B^j \varsigma_{1A} \otimes \varsigma_{jB} + \sum_{i=x,y,z} r_A^i \varsigma_{iA} \otimes \varsigma_{1B} \right.$$

$$\left. + \sum_{i=x,y,z} \sum_{j=x,y,z} r_A^i r_B^j \varsigma_{iA} \otimes \varsigma_{jB} \right)$$

where $r_A^2 \leq 1$ and $r_B^2 \leq 1$ too, and where the equality would hold for *pure* single qubit constituents. Analogous expressions can be constructed for the quaternion and density matrix formalisms by merely replacing varsigmas with sigmas.

We are going to call such state a *simple separable* state. Its individual constituents may not be pure and the resulting biqubit may not be pure either, but it is made of just one pair of well defined separate qubits, pure or not.

Comparing the above expression with Bell states and remembering that varsigmas constitute bases in the fiducial spaces of both qubits, we find the first important difference: all Bell states are of the form

$$\frac{1}{4} \left(\varsigma_{1A} \otimes \varsigma_{1B} \pm \varsigma_{xA} \otimes \varsigma_{xB} \pm \varsigma_{yA} \otimes \varsigma_{yB} \pm \varsigma_{zA} \otimes \varsigma_{zB} \right)$$

There are no $\varsigma_{1A} \otimes \varsigma_{jB}$ and no $\varsigma_{iA} \otimes \varsigma_{1B}$ terms here. This immediately suggests that $r_A = r_B = 0$. And yet $r_A^i r_B^i = \pm 1$. How can this be?

Consider a *mixture* of two simple separable biqubit states:

$$P_\alpha \boldsymbol{p}_{A\alpha} \otimes \boldsymbol{p}_{B\alpha} + P_\beta \boldsymbol{p}_{A\beta} \otimes \boldsymbol{p}_{B\beta}$$

where $P_\alpha + P_\beta = 1$ and they're both positive.

The fiducial representation of the mixture is:

$$\frac{1}{4} \left(\varsigma_{1A} \otimes \varsigma_{1B} \right.$$

$$+ \sum_{i=x,y,z} \left(P_\alpha r_{A\alpha}^i + P_\beta r_{A\beta}^i \right) \varsigma_{iA} \otimes \varsigma_{1B}$$

$$+ \sum_{i=x,y,z} \left(P_\alpha r_{B\alpha}^i + P_\beta r_{B\beta}^i \right) \varsigma_{1A} \otimes \varsigma_{iB}$$

$$\left. + \sum_{i,j=x,y,z} \left(P_\alpha r_{A\alpha}^i r_{B\alpha}^j + P_\beta r_{A\beta}^i r_{B\beta}^j \right) \varsigma_{iA} \otimes \varsigma_{jB} \right)$$

Observe that in general

$$P_\alpha r_{A\alpha}^i r_{B\alpha}^j + P_\beta r_{A\beta}^i r_{B\beta}^j \neq \left(P_\alpha r_{A\alpha}^i + P_\beta r_{A\beta}^i \right) \cdot \left(P_\alpha r_{B\alpha}^j + P_\beta r_{B\beta}^j \right)$$

Therefore if we want to admit a description of mixtures of simple separable biqubit states, each of which may be made of two qubits in some mixed states too, and such a *finite* mixture is called a *separable state*, we must allow the following, more general fiducial representation:

$$
\boldsymbol{\rho}_{AB} = \frac{1}{4}\left(\alpha\,\boldsymbol{\varsigma}_{1A}\otimes\boldsymbol{\varsigma}_{1B} + \sum_{j=x,y,z} r_B^j\boldsymbol{\varsigma}_{1A}\otimes\boldsymbol{\varsigma}_{jB} + \sum_{i=x,y,z} r_A^i\boldsymbol{\varsigma}_{iA}\otimes\boldsymbol{\varsigma}_{1B} \right.
$$
$$
\left. + \sum_{i,j=x,y,z} x_{AB}^{ij}\boldsymbol{\varsigma}_{iA}\otimes\boldsymbol{\varsigma}_{jB} \right) \tag{5.7}
$$

where α may not necessarily be 1 and x_{AB}^{ij} may be independent of r_A^i and r_B^j. There would be nine such x_{AB}^{ij} coefficients, plus three r_A^i coefficients and three r_B^j coefficients, plus the α – altogether sixteen *real* numbers are therefore needed to fully characterize a biqubit state. This number may be reduced to fifteen by imposing a normalization condition such as

$$
p^{\uparrow\uparrow} + p^{\uparrow\downarrow} + p^{\downarrow\uparrow} + p^{\downarrow\downarrow} = 1
$$

which implies that $\alpha = 1$.

It is possible to construct complicated biqubit mixtures made of many components, more than just two, that can get pretty close to an entangled state such as the Bell state $\mid\Psi^-\rangle_{AB}$. This is why it may be sometimes difficult to distinguish between entangled and mixed states experimentally, especially if neither is pure.

Consider, for example, a 50/50 mixture of two simple separable biqubit states described by the following two pairs of vectors:

$$
(\boldsymbol{r}_{A\alpha}, \boldsymbol{r}_{B\alpha}) \quad \text{and} \quad (\boldsymbol{r}_{A\beta}, \boldsymbol{r}_{B\beta})
$$

such that

$$
\boldsymbol{r}_{A\alpha} = -\boldsymbol{r}_{A\beta} \quad \text{and} \quad \boldsymbol{r}_{B\alpha} = -\boldsymbol{r}_{B\beta}
$$

For this mixture we'll get that

$$
P_\alpha r_{A\alpha}^i + P_\beta r_{A\beta}^i = 0
$$
$$
P_\alpha r_{B\alpha}^i + P_\beta r_{B\beta}^i = 0
$$

but

$$
\begin{aligned}
x_{AB}^{ij} &= P_\alpha r_{A\alpha}^i r_{B\alpha}^j + P_\beta r_{A\beta}^i r_{B\beta}^j \\
&= 0.5\left(r_{A\alpha}^i r_{B\alpha}^j + (-r_{A\alpha}^i)(-r_{B\alpha}^j) \right) \\
&= r_{A\alpha}^i r_{B\alpha}^j
\end{aligned}
$$

So here we end up with a state that looks somewhat similar to $\mid \Psi^- \rangle_{AB}$. It's $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$ vanish, but its $\boldsymbol{x}_{AB}$ does not. At the same time though, its $\boldsymbol{x}_{AB}$ is not the same as the one we got for $\mid \Psi^- \rangle_{AB}$. We cannot make a diagonal matrix with none of the diagonal elements vanishing out of $r^i_{A\alpha} r^j_{B\alpha}$.

It turns out that $\boldsymbol{p}_{AB}$ of an entangled state cannot be reproduced by any finite mixture of simple separable biqubit states [30]. But this is not a criterion that is easy to use, especially when we deal with experimental data that is always contaminated with some error.

For a simple separable biqubit state the following trivial observations hold.

If r^i_A and r^j_B do not vanish and $x^{ij}_{AB} = r^i_A r^j_B$ then the biqubit is clearly separable into two individual qubits. If one or both r^i_A and r^j_B vanish, then one or both constituent qubits are fully depolarized. In this case x^{ij}_{AB} must vanish too, if the biqubit is to be separated into two individual qubits.

The separability of x^{ij}_{AB} into r^i_A and r^j_B can be tested easily. The following must hold:

$$
\begin{aligned}
\boldsymbol{r}_A \cdot \boldsymbol{x}_{AB} &= r^2_A \boldsymbol{r}_B, \quad \text{and} \\
\boldsymbol{x}_{AB} \cdot \boldsymbol{r}_B &= \boldsymbol{r}_A r^2_B
\end{aligned}
$$

Another obvious feature of a separable $\boldsymbol{x}_{AB}$ in this context is that

$$
\det \boldsymbol{x}_{AB} = \sum_{i,j,k \in \{x,y,z\}} \epsilon_{ijk} x^{xi} x^{yj} x^{zk} = \sum_{i,j,k \in \{x,y,z\}} \epsilon_{ijk} r^x_A r^i_B r^y_A r^j_B r^z_A r^k_B = 0
$$

because ϵ_{ijk} is fully anti-symmetric whereas $r^i_B r^j_B r^k_B$ is fully symmetric. The vanishing of $\det \boldsymbol{x}_{AB}$ is a *necessary* condition, but not a sufficient one. Still, it is good enough if we just want to check if a given state can at all be separated into two individual qubits.

The above considerations are presented in terms of varsigma coefficients. But when the actual measurements are made, the probabilities that are assembled into a matrix $\boldsymbol{p}^{ij}$ are not varsigma coefficients. They are canonical coefficients instead. How are we to find r^i_A, r^j_B and x^{ij}_{AB}? This can be done easily by contracting $\boldsymbol{p}$ with appropriate combinations of Pauli forms, namely:

$$
\begin{aligned}
r^i_A &= \langle \boldsymbol{\varsigma}^i_A \otimes \boldsymbol{\varsigma}^1_B, \boldsymbol{p}_{AB} \rangle \\
r^j_B &= \langle \boldsymbol{\varsigma}^1_A \otimes \boldsymbol{\varsigma}^j_B, \boldsymbol{p}_{AB} \rangle \\
x^{ij}_{AB} &= \langle \boldsymbol{\varsigma}^i_A \otimes \boldsymbol{\varsigma}^j_B, \boldsymbol{p}_{AB} \rangle
\end{aligned}
$$

where i and j run through x, y and z.

Another way is to replace the canonical vectors in $\boldsymbol{p}_{AB} = \sum_{ij} p^{ij}_{AB} \boldsymbol{e}_{iA} \otimes \boldsymbol{e}_{jB}$ with Pauli vectors using equations (2.12) on page 63, since Pauli vectors are duals of Pauli forms.

5.2   Pauli exclusion principle

But how do we know entangled states exist at all? They can't be made by merely placing two qubits next to each other. This only yields a non-entangled biqubit, i.e., two separate qubits.

The idea that pairs of elementary particles can be entangled goes all the way back to Wolfgang Pauli (1900-1958). Today we know that we can even entangle macroscopic objects, sic!, but back then in 1925 quantum mechanics was strictly a science of atoms and elementary particles. After Niels Bohr (1885-1962) presented his model of the hydrogen atom in 1913 that predicted correctly hydrogen's energy levels, well, at least until physicists had a closer look and found that there was more to it, people turned to other atoms trying to understand their structure. It soon became apparent that even as simple an atom as helium was immensely more complicated than hydrogen. For starters, helium had two separate families of energy levels with seldom observed transitions between one and the other. For a while people even thought that there were two different types of helium configurations. Other atoms' spectra were even more complex.

Eventually Pauli figured out that he could account for some features of alcaline atoms spectra if he (1) postulated that electrons had an additional as yet unrecognized degree of freedom (later called *spin*) and (2) that

> "there can never be two or more equivalent electrons in an atom for which ... the values of all quantum numbers ... are the same. If an electron is present in the atom for which these quantum numbers ... have definite values, this state is *occupied* [42]."

At the time quantum mechanics did not exist in the form we know it today. Werner Karl Heisenberg (1901-1976) was yet to publish his 1925 paper on the "matrix mechanics" and Erwin Schrödinger was yet to publish his 1926 paper on the "wave mechanics". When Paul Dirac (1902-1984) finally merged the matrix and the wave mechanics into quantum mechanics in 1926 and Kronig, Uhlenbeck and Goudsmit identified Pauli's additional degree of freedom as spin it became clear that the way to express Pauli's principle was to antisymmetrize the multielectron wave function. For example, for two electrons A and B whose individual wave functions may be $\mid \Psi \rangle_A$ and $\mid \Phi \rangle_B$ their combined wave function in an atom would have to be

$$\frac{1}{\sqrt{2}} \left(\mid \Psi \rangle_A \otimes \mid \Phi \rangle_B - \mid \Phi \rangle_A \otimes \mid \Psi \rangle_B \right)$$

This way if both functions are identical, i.e., the quantum numbers inside $\mid \Psi \rangle$ and $\mid \Phi \rangle$ are all the same, so that $\mid \Psi \rangle = \mid \Phi \rangle$, the combined wave function is zero. This trick captures the Pauli principle automatically.

In summary, it is the spectra of multielectron atoms that tell us that entangled states such as $\left(\mid \uparrow \rangle_A \otimes \mid \downarrow \rangle_B - \mid \downarrow \rangle_A \otimes \mid \uparrow \rangle_B \right) / \sqrt{2}$ exist.

The Pauli exclusion principle applies to identical particles only, i.e., they must be both of the same type, e.g., two electrons, two neutrons or two protons. They must be fermions too, i.e., their spin must be an odd multiple of $\hbar/2$. If the particles are bosons, which means that their spin is an even multiple of $\hbar/2$, examples of such particles are photons and alpha particles, then their wave functions must be symmetric. This has important macroscopic consequences in terms of statistics and phenomena such as superconductivity and superfluidity.

> In layman terms fermions hate being like the other guys. They're individualistic. If there is a fermion nearby that does something and you're a fermion too, you'll do your best to dress differently, drive a different car, look the other way, and preferably get out of the neighbourhood as soon as an opportunity arises. On the other hand bosons love to be together and to be alike. "I'm having what she's having". If there is a boson driving on a freeway, soon there'll be the whole pack of them driving in the same direction right next to each other – this is how superfluidity works.
>
> Cats are fermions. Dogs are bosons.

Identical particles as such are not very useful for computing. The reason for this is that they cannot be addressed individually. In quantum computing the best results are obtained if qubits are associated with different quantum objects, for example with different nuclei in a molecule. But identical particles can be combined into collective states. Here a quantum state of interest may be associated not with an individual electron, but with a large number of electrons all forced into a single quantum configuration.

But how can electrons be forced into a collective if they are fermions? This can happen only if there is an intermediary agent present. The agent acts like a glue. Electrons in a crystal lattice of niobium, for example, are held together by phonons. Pairs of such phonon-glued electrons, called Cooper pairs, behave like bosons, because their combined spin is an even multiple of $\hbar/2$. Since Cooper pairs are bosons, large collective states are possible. These states are responsible for superconductivity.

Because all elementary particles can be classified either as fermions or as bosons and these can only exist in entangled states, an obvious question is how come we don't normally see entanglement in the macroscopic world around us – apart from

the rather special phenomena of superfluidity and superconductivity, both of which require extremely low temperatures?

The answer to this is that entanglement is technically a superposition. It is a superposition in the biqubit Hilbert space. The biqubit superpositions depolarize just as quickly, if not more so, as single qubit superpositions. The more a biqubit interacts with the environment the faster it depolarizes. And so after a short while the entangled state turns into a mixture of separable biqubit basis states, similarly to what we have seen for a single qubit in the quantronium example.

5.3   A superconducting biqubit

The two quantum electronic devices that are discussed in this section were constructed to demonstrate the entanglement of two qubits.

The first one was made by a group of scientists from the University of Maryland in 2003 [5] and the second one three years later by a group from the University of California, Santa Barbara [50].

A Josephson junction, which is constructed by inserting a very thin insulator between two superconductors, is characterized by the so called *critical current*. The junction is biased by pushing current through it. Up to the critical current, pairs of electrons, the Cooper pairs, flow through the junction unimpeded and without any voltage drop across the junction. This is called the DC Josephson junction regime. But when the current exceeds the critical current, then the voltage drop across the junction suddenly appears and the current begins to oscillate rapidly. This is called the AC Josephson junction regime.

When the junction is biased somewhat below the critical current, i.e., still in the DC Josephson regime, the junction's inductance and capacitance form an anharmonic LC resonator with an anharmonic potential U that can be approximated by a cubic function of position within the junction, as shown in figure 5.1 (A). This is sometimes called a "tilted washboard potential". Strictly speaking U is a function of the *phase* γ of the Cooper pairs wave function within the circuit (they all have the same wave function), but within the junction the phase changes approximately linearly with position and outside of the junction the phase is approximately constant. The left bend of the washboard curve forms a natural potential well, within which discrete energy levels form, and the right bend a natural potential barrier, through which Cooper pairs trapped in the well may tunnel. The height $\Delta U(I)$ of

the potential barrier is a function of the junction bias current and is given by

$$\Delta U(I) = \frac{2\sqrt{2}I_0\Phi_0}{3\pi}\left(1 - \frac{I}{I_0}\right)^{3/2},$$

where I_0 is the critical current and $\Phi_0 = h/2e$, where h is the Planck constant. The potential barrier vanishes when $I \to I_0$. The energy at the bottom of the well corresponds to the classical plasma oscillation frequency $\omega_p(I)$ given by

$$\omega_p(I) = \sqrt{\frac{2\sqrt{2}\pi I_0}{\Phi_0 C}}\left(1 - \frac{I}{I_0}\right)^{1/4},$$

where C is the junction capacitance. The $\omega_p(I)$ also vanishes when $I \to I_0$, but more slowly than $\Delta U(I)$.

The discrete energy levels within the well can be observed by adding small microwave pulses to the bias DC current. If the frequency of the pulse matches the transition frequency between the two levels drawn in the diagram, Cooper pairs trapped within the lower level absorb the energy and jump to the upper level, which is characterized by a large tunneling rate. And so they then tunnel through the barrier. This is illustrated by the right-pointing arrow in figure 5.1 (A). After the tunneling event the junction becomes momentarily depleted of carriers and behaves like an open switch across which a macroscopically measurable voltage develops.

Such qubits, which are made of collective excitations of up to 10^9 paired electrons, depending on the size of the junction, are called *phase qubits*. Recent advances in the phase qubit technology made it possible to carry out their full characterization. The group from Santa Barbara was able to view the full traversal of the qubit, across the Bloch ball during a Ramsey measurement [51]. Like the quantronium circuit, phase qubits are macroscopic devices that behave quantum-mechanically. Also, like the quantronium circuit, they have to be immersed in a cryo-bath, close to the absolute zero, to work.

But let us return to the Maryland group biqubit shown in Figure 5.1 (B). Here we have two phase qubits characterized by their critical currents I_{c1} and I_{c2} of $14.779\,\mu\text{A}$ and $15.421\,\mu\text{A}$ respectively, both shunted by capacitors C_j of $4.8\,\text{pF}$ each, which help stabilize the qubits, and coupled through the capacitor C_c of $0.7\,\text{pF}$.
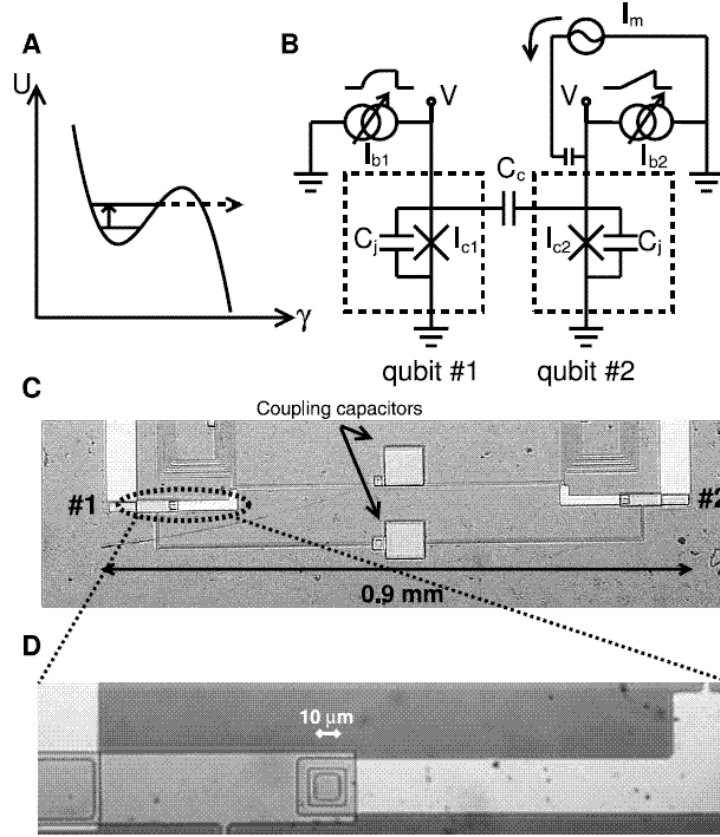
Figure 5.1: A Josephson junction biqubit. Figure (A) illustrates how a two-level quantum system, i.e., a qubit, forms within the junction when it is biased in a special way. Figure (B) shows the schematic diagram of the device. Figures (C) and (D) show the photographs of the device.

Qubit #1, i.e., the one on the left, is DC biased and the bias current that flows through it, I_{b1}, is $14.630\,\mu$A. Qubit #2, the one on the right, is biased with a linear ramp that in effect allows for repetitive scanning of the biqubit parameters.

Figure 5.1 (C) shows the photograph of the actual device. There are two coupling capacitors in the device visible in the centre of the photograph. The role of the

lower capacitor is to short out parasitic inductance in the ground line. It forms the effective C_c together with the upper capacitor. The Josephson junctions are inside the two narrow horizontal features on both sides of the photograph, one of which is shown in magnification in figure 5.1 (D). Here we can see two strips made of niobium that overlap in a square box a little to the left of the center of figure D. This is the junction itself. The two overlapping niobium strips are separated by a thin layer of aluminum oxide. Each Josephson junction is quite large, the side of the box being $10\,\mu$m long. The distance between the two junctions is $0.7\,$mm - which is *huge* by quantum mechanics standards. It is a macroscopic distance.

The biqubit is observed in a way that is very similar to a single phase qubit observation technique. It has a certain characteristic and discrete energy spectrum. A microwave signal I_m (see figure 5.1 (B)) is applied through the bias lines to the right qubit directly, but through the coupling capacitor to the left qubit too—so in effect it is the whole biqubit that is irradiated. This induces transitions from the ground state of the biqubit to its higher energy states, but only if the applied microwave signal has its frequency matching the energy gap. Higher energy states are closer to the knees of the washboard potentials for both qubits, and so they have a higher probability of escaping from the potential wells of both junctions. This creates an open circuit condition, because the junctions run out of carriers and this, in turn, manifests as surges of DC voltage V in the right hand side of the circuit and can be detected easily.

Ramping I_{b2} for a microwave signal of fixed frequency has the effect of changing the parameters of the biqubit so that eventually we come across the ones for which the energy absorption takes place.

This is illustrated in figure 5.2 (A). Here the microwave frequency f is set to $4.7\,$GHz. Δ is the observed escape rate normalized against the escape rate measured in the absence of the microwave agitation, i.e., $\Delta = (\Gamma_m - \Gamma)/\Gamma$, where Γ_m is the microwave induced escape rate and Γ is the escape rate in the absence of the microwave signal. The normalized escape rate is plotted against the bias current I_{b2}.

Let us call the lower energy state $\mid 0\rangle_{1,2}$ and the higher energy state $\mid 1\rangle_{1,2}$ for qubits #1 and #2 respectively.

There is a well defined Lorentzian absorption peak in Figure 5.2 (A) that corresponds to a transition

$$\mid 0\rangle_1 \otimes \mid 0\rangle_2 \to \frac{1}{\sqrt{2}}\left(\mid 0\rangle_1 \otimes \mid 1\rangle_2 - \mid 1\rangle_1 \otimes \mid 0\rangle_2\right).$$
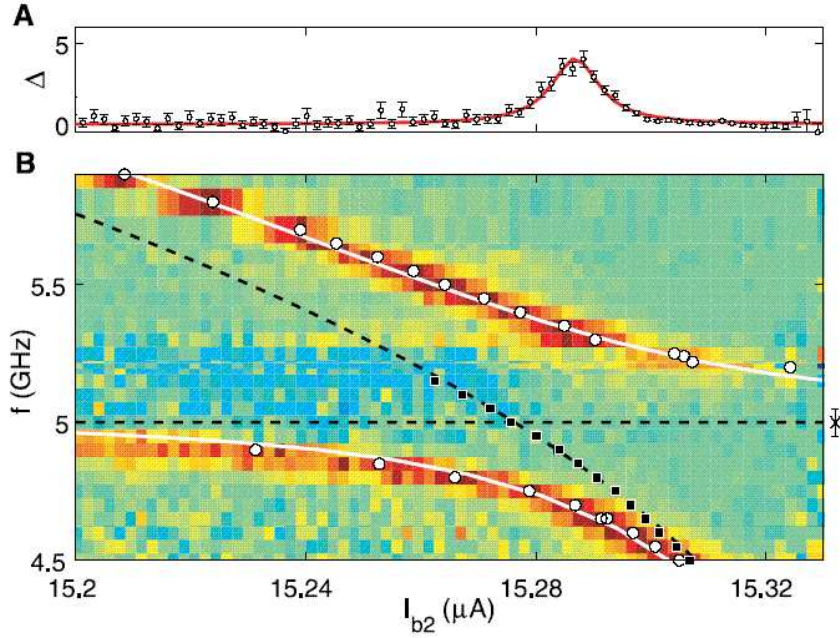
Figure 5.2: Figure (A) – absorption spectrum of the biqubit for a fixed microwave frequency $f = 4.7$ GHz in function of the Josephson junction bias current I_{b2}. Figure (B) – a two-dimensional map obtained by varying both the microwave frequency f and the Josephson junction bias current I_{b2}.

By varying both f and I_{b2} and repeating the measurement for each point in the (f, I_{b2}) plane up to 100,000 times, sic!, the researchers arrived at the histogram map shown in figure 5.2 (B). The colour in the histogram corresponds to the normalized escape rate: red means high, blue means low. For a given value of I_{b2} there are two absorption peaks that correspond to transitions:

$$| \, 0 \rangle_1 \otimes | \, 0 \rangle_2 \to \frac{1}{\sqrt{2}} \left(| \, 0 \rangle_1 \otimes | \, 1 \rangle_2 \pm | \, 1 \rangle_1 \otimes | \, 0 \rangle_2 \right).$$

The tiny white circles mark the exact measured locations of the absorption peaks. Solid white lines mark theoretically predicted[3] locations of the absorption peaks for

---

[3]The theoretical analysis of the biqubit and the resulting computations are quite non-trivial and beyond the scope of this text. Interested reader will find more details in [5].

the transitions to the entangled states investigated here. The black dashed lines correspond to transitions between the ground state $\mid 0\rangle_1 \otimes \mid 0\rangle_2$ and the two *un-entangled* states $\mid 0\rangle_1 \otimes \mid 1\rangle_2$ and $\mid 1\rangle_1 \otimes \mid 0\rangle_2$.

The agreement between the measured data and the theoretical predictions based on the assumption that the transitions are from the ground state to one of the two *entangled* states is quite exceptional.

The measured dependency of the transition rate on the bias current I_{b2} can be Fourier transformed into the frequency space. The width of the absorption peak in this space yields the lifetime of the entangled state, which turns out to be about 2 ns. After this time, the state decays into a mixture of separable states.

What is so remarkable about this beautiful experiment (and device) is that not only does it demonstrate the existence of entangled states, but it entangles two heavy *macroscopic* objects separated by a large *macroscopic* distance of 0.7 mm. It is the cooling to the near absolute zero that makes it possible for the entangled state of the two phase qubits to stretch this far. Cooling freezes off interactions with the environment that would otherwise destroy both the entanglement and the qubits themselves.

We see here again that entangled states and quantum behavior are not restricted to very small objects such as elementary particles and to very small distances such as encountered in the interiors of alcaline atoms.

On the other hand, if the only evidence in favour of the existence of entangled states were atomic or electronic device spectra we might be justified in holding back our enthusiasm. After all, one could perhaps come up with another theory that would reproduce the observed spectra in some other way. To be truly convinced that entanglement is not just an artifact of quantum calculus but a physical phenomenon we need more evidence. We need to construct an entangled state and then perform a full set of measurements on it so as to reproduce probability matrices such as the ones derived for the Bell states in section 5.1.

This has been done for pairs of protons and pairs of photons, the latter separated by a distance as large as 600 m [2]. It has been done even for two macroscopic samples of caesium gas each comprising 10^{12} atoms [34]. And the Santa Barbara group did this for a phase biqubit, but it took full three years of technology and methodology improvements after the Maryland biqubit demonstration.

A full set of quantum state measurements that reproduces its whole probability or density matrix is called quantum state tomography. It is a difficult measurement. Recall that in our quantronium example we only saw the r^z component of the polarization vector $\boldsymbol{r}$. But here we need to measure r^x, r^y and r^z for both qubits as well as correlation coefficients x^{ij}.

The Santa Barbara researchers demonstrated full quantum state tomography for a single qubit first, and for a biqubit. The devices they used in their experiments are shown in Figure 5.3.
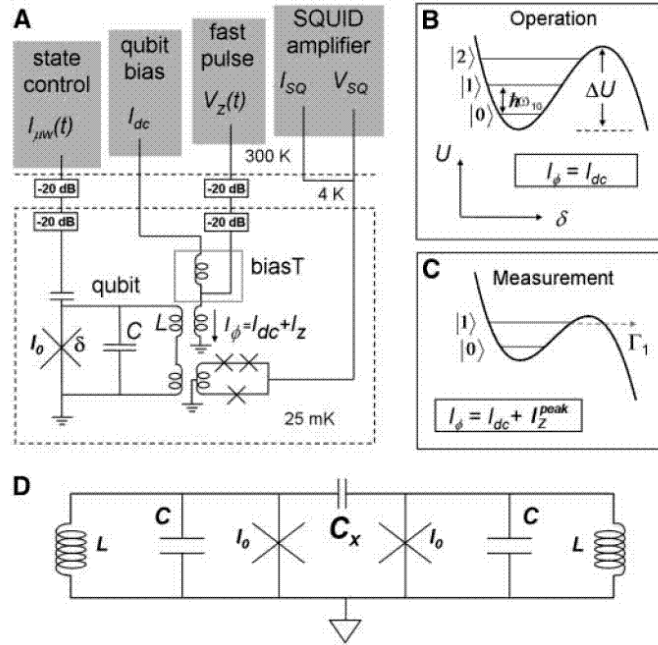

Figure 5.3: A schematic diagram of the Santa Barbara biqubit. (A) a single qubit diagram that shows qubit manipulation and measurement circuitry, (B) and (C) the LC resonator potential during qubit manipulation and measurement, (D) a simplified biqubit diagram that shows the coupling capacitor C_x.

A distinguishing feature of this device is qubit isolation. We no longer apply the bias, the measuring pulse, and the microwave signal directly to the qubit's circuit. Instead, each qubit is manipulated and measured through the loop inductance of $L = 850\,\text{pH}$. The shunting capacitors that stabilize the qubits both have the capacitance $C = 1.3\,\text{pF}$ and the coupling capacitor is $C_x = 3\,\text{fF}$. Amorphous silicon nitride is used as a dielectric in the shunting capacitors, because its loss tangent is very small, of the order of 10^{-4}, which yields a fairly long energy relaxation time of about 170 ns.

When operating normaly both qubits are biased as shown in Figure 5.3 (B), but they are conditioned so that only the two bottom energy levels, $|\,0\,\rangle$ and $|\,1\,\rangle$ are filled. When the qubits are measured, a strong current pulse I_z is applied that changes the junction bias as shown in Figure 5.3 (C). The $|\,1\,\rangle$ state is then flushed out of the cubic well and this event is picked up by the on-chip SQUID amplifier.

> A SQUID, the acronym stands for Superconducting Quantum Interference Device, is another extremely sensitive Josephson junction device that is used to measure changes in the magnetic field flux.

Both qubits are biased so that the transition between each qubit's $|\,0\,\rangle$ and $|\,1\,\rangle$ states occurs at $\omega_{10} = 2\pi \times 5.1\,\text{GHz}$. An experiment begins by freezing and waiting—both qubits drop naturally to $|\,0\,\rangle$. Section 5.11.4, page 260, will explain in more detail how this happens, but the intuitive understanding that when things are left on their own in a cool place they calm down, is just fine at this stage. Now we flip qubit #2 by sending it an appropriate Rabi pulse of $10\,\text{ns}$ duration and the biqubit ends up in the $|\,0\,\rangle \otimes |\,1\,\rangle$ state. Because this state is not an eigenstate of the biqubit Hamiltonian, the biqubit evolves as follows

$$|\,\Psi_{12}(t)\,\rangle = \cos\left(\frac{St}{2\hbar}\right)|\,0\,\rangle \otimes |\,1\,\rangle - i\sin\left(\frac{St}{2\hbar}\right)|\,1\,\rangle \otimes |\,0\,\rangle,$$

where $S/h = 10\,\text{MHz}$. So we don't have to do anything to rotate the biqubit at this stage other than wait a certain t. Having waited a given t we can measure the qubit and by repeating the experiment 1,000 times we can arrive at the probabilities P_{00}, P_{01}, P_{10} and P_{11}. The observed probabilities are consistent with the idea that the biqubit becomes entangled after about $16\,\text{ns}$ forming a state

$$\frac{1}{\sqrt{2}}\left(|\,0\,\rangle \otimes |\,1\,\rangle - i\,|\,1\,\rangle \otimes |\,0\,\rangle\right).$$

Waiting while a qubit, or a biqubit performs a natural (e.g., Larmor) rotation is an often deployed trick in quantum computing and as such constitutes a gate.

In order to fully diagnose the state and make sure that the qubit is indeed entangled we have to measure probabilities of finding the qubits in states such as $|\rightarrow\rangle$ and $|\otimes\rangle$ as well. In other words, we must perform the full tomography of the biqubit state.

We do this by subjecting the biqubit to a yet another Rabi pulse that rotates its individual qubits by $90°$ about the $\boldsymbol{e}_x$ or $\boldsymbol{e}_y$ directions prior to the measurement.

Repeating biqubit preparation and measurement procedures 20,000 times for each combination of directions we obtain probabilities $P_{\uparrow\uparrow}$, $P_{\uparrow\rightarrow}$, and similar and then assemble them into a density matrix shown in Figure 5.4.
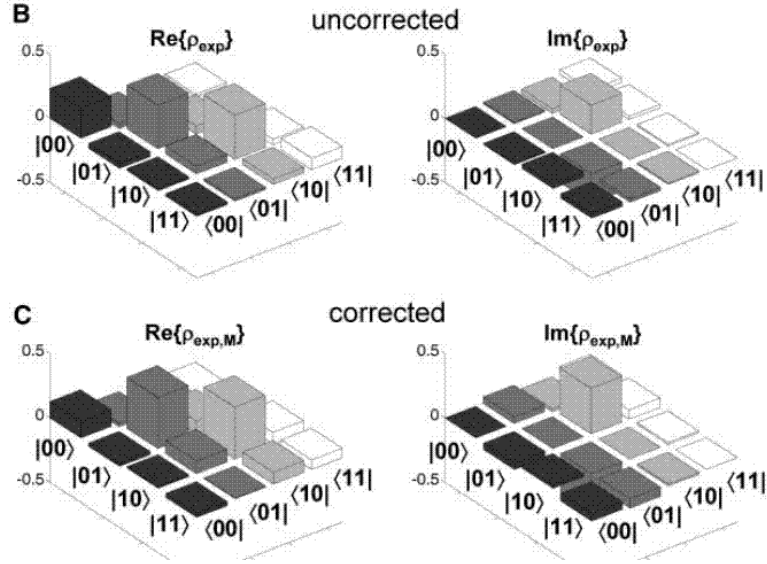
Figure 5.4: State tomography of the Santa Barbara biqubit.

The upper pannel of Figure 5.4 shows the results obtained from the raw proba-
bility data, and the lower pannel shows the density matrix "corrected" for known
inefficiencies of single qubit measurements.

The expression

$$F = \text{Tr}\sqrt{\boldsymbol{\sigma}^{1/2}\boldsymbol{\rho}_{\text{exp}}\boldsymbol{\sigma}^{1/2}}$$

where $\boldsymbol{\sigma}$ is the theoretically expected density matrix and $\boldsymbol{\rho}_{\text{exp}}$ is the experimentally
measured one provides us with a convenient estimate of the combined accuracy of
the state preparation and its tomography in the form of a single number. This
number is called *fidelity* of the reconstructed (from the measurements) quantum
state. If $\boldsymbol{\rho}_{\text{exp}} = \boldsymbol{\sigma}$ the fidelity is 100%.

The fidelity of the state reconstructed in Figure 5.4 (B), i.e., in the upper pan-
nel, is 75% and the fidelity of the state corrected for the known measurement
inefficiencies of single qubits, shown in the lower pannel, is 87%.

An in-depth theoretical analysis of the device that takes various environmental
effects and known characteristics of the device into account shows that the fidelity

should be 89%. This leaves only 3% of the *infidelity* unaccounted for, which is a very impressive quantum device modeling result.

Is the observed state indeed entangled? An inspection of the density matrix suggests so. The imaginary components $\mid 01\rangle\langle 10 \mid$ and $\mid 10\rangle\langle 01 \mid$ have almost the same value as real components $\mid 01\rangle\langle 01 \mid$ and $\mid 10\rangle\langle 10 \mid$, which is what we would expect for the state $(\mid 0\rangle\otimes\mid 1\rangle - i\mid 1\rangle\otimes\mid 0\rangle)/\sqrt{2}$.

But isn't it possible that a mixture could be constructed that would get pretty close to the observed density matrix? Is there a way to demonstrate unequivocally that the observed state is indeed entangled by some well posed criterion other than just looking at and comparing the bars on the graph? This, as it turns out, is a non-trivial question, the answer to which was found only in 1996. We are going to discuss the solution to this problem in Section 5.10, page 234.

5.4   An atom and a photon

Demonstrating entanglement with elaborate quantum electronic circuits even though of obvious practical interest may leave one pondering if a shortcircuit or some specific circuit design feature is not responsible for the observed correlations rather than fundamental physics. After all, these are complicated devices. Their fabrication is difficult and their operation complex. This is, of course, a sentiment, because a great deal of tests, checks and theoretical analysis goes into the design of the device and the experiment. Nevertheless, yielding to this sentiment we may ask if entanglement can be demonstrated using just two atoms or just two elementary particles? Such a demonstration would, at least in principle, prove that the observed behaviour reflects the law of nature and is not an electronic artifact.

Numerous experiments of this type have been performed and are still being performed today. Photons are especially suitable, because they are relatively immune to environmental decoherence—a photon may travel almost undisturbed across the whole observable universe, to be registered by an astronomer's telescope, still with sufficient information content to let us make inferences about its source.

The experiment discussed in this section was carried out by Blinov, Moehring, Duan and Monroe from the University of Michigan in 2004 [7]. It is perhaps one of the cleanest and most elegant demonstrations of entanglement. In this experiment it is a single atom and a single photon emitted by the atom that are entangled and measured.
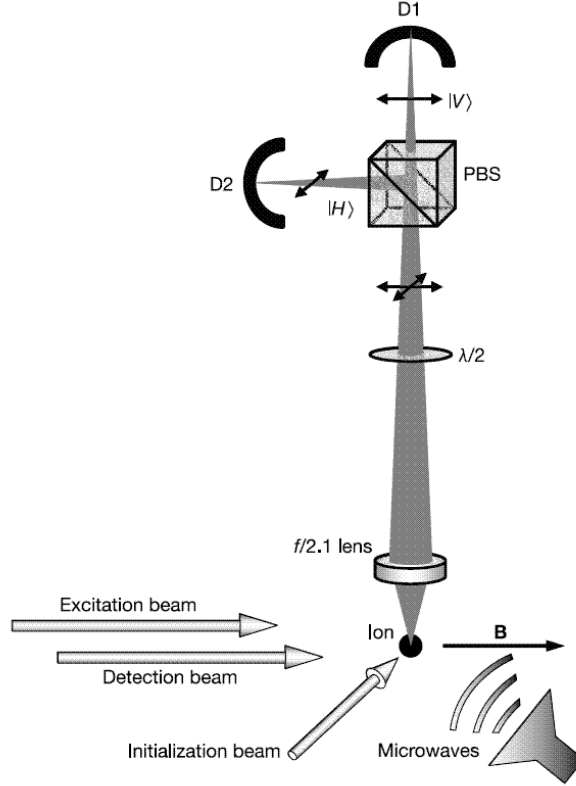
Figure 5.5: A schematic diagram of an apparatus used in the atom-photon entanglement experiment.

Figure 5.5 shows a schematic diagram of an apparatus used in the measurement. A single positively charged ion of cadmium, $^{111}\mathrm{Cd}^+$, is held in an asymmetric-quadrupole radio frequency trap about 0.7 mm across, to which a magnetic field of approximately 0.7 Gauss is applied in order to provide the $\boldsymbol{e}_z$ direction. The ion is manipulated by a combination of optical and microwave pulses. In response to these manipulations the ion emits a single photon, which is collected by the $f/2.1$ imaging lens and directed towards the $\lambda/2$ waveplate. The waveplate is used to rotate the photon polarization, which in this setting is like switching from $p^{\uparrow\mathrm{atom}\,\uparrow\mathrm{photon}}$ to, say, $p^{\uparrow\mathrm{atom}\,\rightarrow\mathrm{photon}}$. The photon is then directed towards a polarizing beamsplitter, marked as "PBS" in the diagram, and then sent towards one of the two photon-

counting photomultiplier tubes (PMTs) that can detect a single photon with about 20% efficiency. The PMT detector D1 is set up to detect photons polarized in the plane of the figure, these are called $| V \rangle$-photons, and the other detector, D2, is set up to detect photons polarized in the plane perpendicular to the plane of the figure, these are called $| H \rangle$-photons. The $| V \rangle$ and $| H \rangle$ states of the photon are like qubit states $| 0 \rangle$ and $| 1 \rangle$.

The purpose of the experiment is to demonstrate the quantum entanglement between the ion and the photon emitted by it. Whereas the state of the photon is measured by the PMT detectors shown on top of the diagram, reading the state of the ion is performed with a specially polarized $200\,\mu$s optical detection pulse beamed at the ion. The ion responds to the pulse by fluorescing differently depending on its state. Here the ion qubit read-out efficiency is greater than 95%.

Prior to the read-out the ion's quantum state can be subjected to a Rabi rotation by irradiating it with a microwave pulse. So this way we can measure, say, $p^{\uparrow \text{atom} \uparrow \text{photon}}$ and $p^{\rightarrow \text{atom} \uparrow \text{photon}}$ as well.

The sequence of operations the ion is subjected to is roughly as follows.

First the ion is initialized in the $|\uparrow\rangle$ state by a combination of a $30\,\mu$s polarized optical pulse and a $15\,\mu$s microwave rotation.

It is then excited to a short lived higher energy state called $^2\mathrm{P}_{3/2} \, | \, 2, 1 \rangle$ by a $50\,$ns polarized optical pulse. The $^2\mathrm{P}_{3/2} \, | \, 2, 1 \rangle$ state decays after about $3\,$ns either back to $|\uparrow\rangle$ or to a state with a somewhat higher energy here called $|\downarrow\rangle$. The decay of $^2\mathrm{P}_{3/2} \, | \, 2, 1 \rangle$ to $|\uparrow\rangle$ is accompanied by emission of an $| H \rangle$ photon and the decay of $^2\mathrm{P}_{3/2} \, | \, 2, 1 \rangle$ to $|\downarrow\rangle$ is accompanied by emission of a $| V \rangle$ photon:

$$
\begin{aligned}
^2\mathrm{P}_{3/2} \, | \, 2, 1 \rangle &\rightarrow |\uparrow\rangle + | H \rangle \\
^2\mathrm{P}_{3/2} \, | \, 2, 1 \rangle &\rightarrow |\downarrow\rangle + | V \rangle
\end{aligned}
$$

The energy gap that separates $|\uparrow\rangle$ and $|\downarrow\rangle$ is about $1\,$MHz.

After the initial preparation procedure the ion is allowed to rest for about $1\,\mu$s and then it is irradiated again with another $15\,\mu$s microwave rotation pulse. Finally the ion is irradiated with a $200\,\mu$s polarized optical detection pulse that lets us read the ion resident qubit.

Theoretical analysis of this process reveals that the ion and the emitted photon must be entangled and the resulting state is

$$
(| H \rangle \otimes |\uparrow\rangle + 2 \, | V \rangle \otimes |\downarrow\rangle) / \sqrt{3}
$$

Because the photon detector is markedly less efficient than the ion qubit detector the experimenters measure probabilities of detecting an atomic qubit state $|\uparrow\rangle$ or

$|\!\downarrow\rangle$ *conditioned* upon detecting an emitted photon either in the $|H\rangle$ or in the $|V\rangle$ states, given 1,000 successful trials.
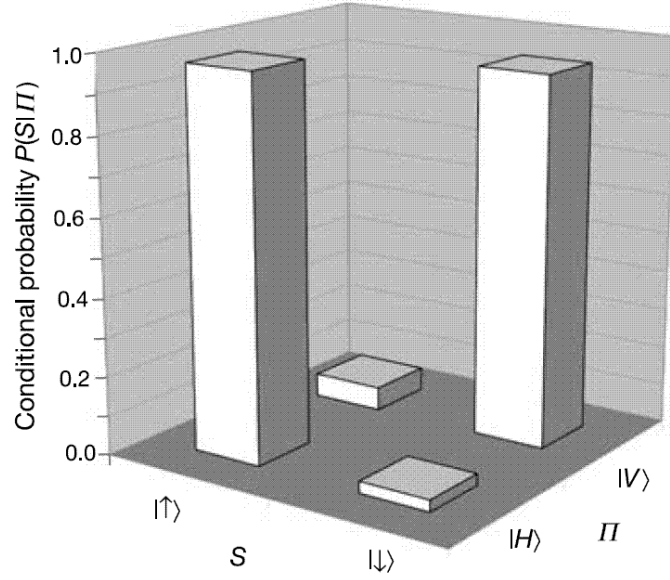

Figure 5.6: Measured conditional probabilities in the original basis – no atomic or photonic qubit rotation before the measurement.

The results for the original basis, i.e., no atomic or photonic rotation before the measurement, are shown in figure 5.6. Here we find that

$$\left(\begin{array}{cc} p^{\uparrow,H} & p^{\uparrow,V} \\ p^{\downarrow,H} & p^{\downarrow,V} \end{array} \right) \equiv \left(\begin{array}{cc} p^{\uparrow\uparrow} & p^{\uparrow\downarrow} \\ p^{\downarrow\uparrow} & p^{\downarrow\downarrow} \end{array} \right) = \left(\begin{array}{cc} 0.97 \pm 0.01 & 0.06 \pm 0.01 \\ 0.03 \pm 0.01 & 0.94 \pm 0.01 \end{array} \right)$$

Now the experimenters rotate the $\lambda/2$ waveplate and the atomic qubit (by applying a microwave pulse after the emission) so as to rotate both through a Bloch angle of 90° eventually. The rotation of the atom-resident qubit is not clean though. It is loaded with an additional phase factor due to the phase of the microwave signal – this is really the angle that the atomic qubit subtends with the photonic qubit in the $\boldsymbol{e}_x \times \boldsymbol{e}_y$ plane. This angle can be adjusted and varying it results in the correlation fringes shown in figure 5.7 (a). The fringes correspond to $p^{\rightarrow\rightarrow}$ and $p^{\rightarrow\leftarrow}$. By locking ourselves on the point of highest correlation in figure 5.7 (a) we can finally arrive at the results shown in figure 5.7 (b).
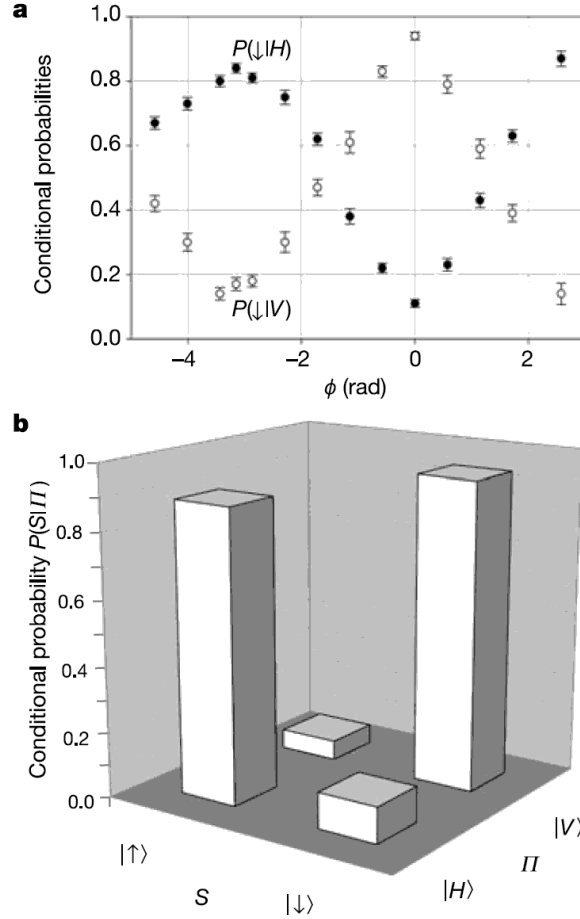
Figure 5.7: Measured conditional probabilities after a Bloch rotation of both qubits by $90°$.

The actual probabilities are

$$
\begin{pmatrix} p^{\rightarrow\rightarrow} & p^{\rightarrow\leftarrow} \\ p^{\leftarrow\rightarrow} & p^{\leftarrow\leftarrow} \end{pmatrix} = \begin{pmatrix} 0.89 \pm 0.01 & 0.06 \pm 0.01 \\ 0.11 \pm 0.01 & 0.94 \pm 0.01 \end{pmatrix}
$$

This is not full tomography as we have seen done for the Santa Barbara qubit, but we can clearly observe correlations. Both qubits appear aligned for both mea-

surement angles, whereas the probability of finding the qubits counter-aligned is very low in both cases.

5.5   A biqubit in a rotated frame

Consider the Bell state $\mid \Psi^- \rangle_{AB}$ for which

$$\boldsymbol{p}_{AB} = \frac{1}{4} \left(\boldsymbol{\varsigma}_{1A} \otimes \boldsymbol{\varsigma}_{1B} - \boldsymbol{\varsigma}_{xA} \otimes \boldsymbol{\varsigma}_{xB} - \boldsymbol{\varsigma}_{yA} \otimes \boldsymbol{\varsigma}_{yB} - \boldsymbol{\varsigma}_{zA} \otimes \boldsymbol{\varsigma}_{zB} \right)$$

We saw in section 4.5, page 130 that Pauli vectors $\boldsymbol{\varsigma}_i$ transform like normal 3D vectors under rotations. So if we were to rotate the polarization filter of qubit A in the $\boldsymbol{e}_x \times \boldsymbol{e}_z$ plane by angle θ_A we'd find that

$$\begin{aligned}
\boldsymbol{\varsigma}_{xA} &= \cos\theta_A \boldsymbol{\varsigma}_{x'A} - \sin\theta_A \boldsymbol{\varsigma}_{z'A} \\
\boldsymbol{\varsigma}_{zA} &= \sin\theta_A \boldsymbol{\varsigma}_{x'A} + \cos\theta_A \boldsymbol{\varsigma}_{z'A}
\end{aligned}$$

and similarly if we were to rotate the polarization filter of qubit B in the same plane by angle θ_B we'd find that

$$\begin{aligned}
\boldsymbol{\varsigma}_{xB} &= \cos\theta_B \boldsymbol{\varsigma}_{x'B} - \sin\theta_B \boldsymbol{\varsigma}_{z'B} \\
\boldsymbol{\varsigma}_{zB} &= \sin\theta_B \boldsymbol{\varsigma}_{x'B} + \cos\theta_B \boldsymbol{\varsigma}_{z'B}
\end{aligned}$$

What are going to be the probability readings for the biqubit now? We can find $\boldsymbol{p}'_{AB}$ by rotating its component varsigmas and remembering that:

$$\begin{aligned}
\boldsymbol{\varsigma}_{1A} &= \boldsymbol{\varsigma}_{1'A} \\
\boldsymbol{\varsigma}_{yA} &= \boldsymbol{\varsigma}_{y'A} \\
\boldsymbol{\varsigma}_{1B} &= \boldsymbol{\varsigma}_{1'B} \\
\boldsymbol{\varsigma}_{yB} &= \boldsymbol{\varsigma}_{y'B}
\end{aligned}$$

This yields

$$\begin{aligned}
\boldsymbol{p}'_{AB} = \; & \frac{1}{4} \Big(\boldsymbol{\varsigma}_{1'A} \otimes \boldsymbol{\varsigma}_{1'B} - \boldsymbol{\varsigma}_{y'A} \otimes \boldsymbol{\varsigma}_{y'B} \\
& - (\cos\theta_A \boldsymbol{\varsigma}_{x'A} - \sin\theta_A \boldsymbol{\varsigma}_{z'A}) \otimes (\cos\theta_B \boldsymbol{\varsigma}_{x'B} - \sin\theta_B \boldsymbol{\varsigma}_{z'B}) \\
& - (\sin\theta_A \boldsymbol{\varsigma}_{x'A} + \cos\theta_A \boldsymbol{\varsigma}_{z'A}) \otimes (\sin\theta_B \boldsymbol{\varsigma}_{x'B} + \cos\theta_B \boldsymbol{\varsigma}_{z'B}) \Big)
\end{aligned}$$

Let us gather terms that multiply $\varsigma_{x'A} \otimes \varsigma_{x'B}$, $\varsigma_{x'A} \otimes \varsigma_{z'B}$, $\varsigma_{z'A} \otimes \varsigma_{x'B}$ and $\varsigma_{z'A} \otimes \varsigma_{z'B}$:

$$
\begin{aligned}
\boldsymbol{p}'_{AB} \;=\; &\frac{1}{4}\Big(\varsigma_{1'A} \otimes \varsigma_{1'B} - \varsigma_{y'A} \otimes \varsigma_{y'B} \\
&- (\cos\theta_A \cos\theta_B + \sin\theta_A \sin\theta_B)\,\varsigma_{x'A} \otimes \varsigma_{x'B} \\
&- (-\cos\theta_A \sin\theta_B + \sin\theta_A \cos\theta_B)\,\varsigma_{x'A} \otimes \varsigma_{z'B} \\
&- (-\sin\theta_A \cos\theta_B + \cos\theta_A \sin\theta_B)\,\varsigma_{z'A} \otimes \varsigma_{x'B} \\
&- (\sin\theta_A \sin\theta_B + \cos\theta_A \cos\theta_B)\,\varsigma_{z'A} \otimes \varsigma_{z'B}\Big)
\end{aligned}
$$

Recal that

$$
\begin{aligned}
\sin(\theta_A - \theta_B) &= \sin\theta_A \cos\theta_B - \cos\theta_A \sin\theta_B \\
\cos(\theta_A - \theta_B) &= \cos\theta_A \cos\theta_B + \sin\theta_A \sin\theta_B
\end{aligned}
$$

consequently

$$
\begin{aligned}
\boldsymbol{p}'_{AB} \;=\; &\frac{1}{4}\Big(\varsigma_{1'A} \otimes \varsigma_{1'B} - \varsigma_{y'A} \otimes \varsigma_{y'B} \\
&- \cos(\theta_A - \theta_B)\,\varsigma_{x'A} \otimes \varsigma_{x'B} \\
&- \sin(\theta_A - \theta_B)\,\varsigma_{x'A} \otimes \varsigma_{z'B} \\
&+ \sin(\theta_A - \theta_B)\,\varsigma_{z'A} \otimes \varsigma_{x'B} \\
&- \cos(\theta_A - \theta_B)\,\varsigma_{z'A} \otimes \varsigma_{z'B}\Big)
\end{aligned}
$$

In order to extract the actual probabilities we need to switch from the rotated Pauli vectors $\varsigma_{i'}$ to the rotated canonical basis vectors $\boldsymbol{e}_{\alpha'}$, $\alpha' = 0, 1, 2, 3$, as we did in section 5.1. This leads to the following mappings between $\varsigma_{i'A} \otimes \varsigma_{j'B}$ and probability matrices:

$$
\varsigma_{1'A} \otimes \varsigma_{1'B} \;=\; \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}
$$

$$
\varsigma_{x'A} \otimes \varsigma_{x'B} \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

$$
\varsigma_{y'A} \otimes \varsigma_{y'B} \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

$$\varsigma_{z'A} \otimes \varsigma_{z'B} \;=\; \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Additionally we need to find matrices for $\varsigma_{x'A} \otimes \varsigma_{z'B}$ and $\varsigma_{z'A} \otimes \varsigma_{x'B}$. Since $\varsigma_{x'A,B} = e_{2'A,B}$ and $\varsigma_{z'A,B} = e_{0'A,B} - e_{1'A,B}$ it is easy to see that

$$\varsigma_{x'A} \otimes \varsigma_{z'B} \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\varsigma_{z'A} \otimes \varsigma_{x'B} \;=\; \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Combining it all together yields

$$
\begin{aligned}
p'_{AB} \;=\; \frac{1}{4}\Bigg(&\left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) \\[2mm]
&- \cos\left(\theta_A - \theta_B\right) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\[2mm]
&- \cos\left(\theta_A - \theta_B\right) \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\[2mm]
&- \sin\left(\theta_A - \theta_B\right) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\[2mm]
&+ \sin\left(\theta_A - \theta_B\right) \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Bigg)
\end{aligned}
$$

We could just add it all up now, but it is convenient to make use of the following trigonometric identities

$$\frac{1 - \cos\left(\theta_A - \theta_B\right)}{2} = \sin^2\frac{\theta_A - \theta_B}{2}$$
$$\frac{1 + \cos\left(\theta_A - \theta_B\right)}{2} = \cos^2\frac{\theta_A - \theta_B}{2}$$

Similar formulae for $1 \pm \sin\left(\theta_A - \theta_B\right)$ are somewhat clumsier because we end up with a $90°$ angle thrown in:

$$\frac{1 - \sin\left(\theta_A - \theta_B\right)}{2} = \cos^2\frac{\theta_A - \theta_B + 90°}{2}$$
$$\frac{1 + \sin\left(\theta_A - \theta_B\right)}{2} = \sin^2\frac{\theta_A - \theta_B + 90°}{2}$$

Let $\theta_A - \theta_B = \theta_{AB}$ for short, then at long last

$$\boldsymbol{p}'_{AB} = \frac{1}{2}\begin{pmatrix} \sin^2\frac{\theta_{AB}}{2} & \cos^2\frac{\theta_{AB}}{2} & \sin^2\frac{\theta_{AB}+90°}{2} & \frac{1}{2} \\ \cos^2\frac{\theta_{AB}}{2} & \sin^2\frac{\theta_{AB}}{2} & \cos^2\frac{\theta_{AB}+90°}{2} & \frac{1}{2} \\ \cos^2\frac{\theta_{AB}+90°}{2} & \sin^2\frac{\theta_{AB}+90°}{2} & \sin^2\frac{\theta_{AB}}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \tag{5.8}$$

Observe that for $\theta_A = \theta_B$ so that $\theta_{AB} = 0$ the probability matrix does not change at all, i.e., for the Bell state $\mid \Psi^-\rangle_{AB}$ the probability of finding both qubits aligned in *any* direction, not just the original $\boldsymbol{e}_x$, $\boldsymbol{e}_y$ and $\boldsymbol{e}_z$ is zero.

But evaluating probabilities at angles other than $\theta_{AB} = 0$ or $\theta_{AB} = 90°$ (which is what basically sits in terms such as $p^{\uparrow\rightarrow}$) reveals something quite peculiar. Let us evaluate

$$p^{\uparrow\nearrow} + p^{\nearrow\rightarrow} - p^{\uparrow\rightarrow}$$

where $\nearrow$ stands for a polarization axis that is tilted by $45°$. This can be rewritten as

$$p^{\uparrow\uparrow}(0°, 45°) + p^{\uparrow\uparrow}(45°, 90°) - p^{\uparrow\uparrow}(0°, 90°)$$

where the first angle in the bracket is θ_A and the second angle is θ_B, or in terms of θ_{AB}:

$$p^{\uparrow\uparrow}(45°) + p^{\uparrow\uparrow}(45°) - p^{\uparrow\uparrow}(90°)$$
$$= \frac{1}{2}\sin^2\frac{45°}{2} + \frac{1}{2}\sin^2\frac{45°}{2} - \frac{1}{2}\sin^2\frac{90°}{2}$$
$$= \sin^2(\pi/8) - 0.5 \cdot \sin^2(\pi/4) = -0.10355 \tag{5.9}$$

So what?

The fact that this number is negative tells us an insightful story about the world of quantum physics.

5.6   Bell inequality

Let us consider

$$p^{\uparrow_A, \nearrow_B} + p^{\nearrow_A, \rightarrow_B} \tag{5.10}$$

for the Bell state $\mid \Psi^- \rangle_{AB}$. We have marked the arrows clearly with A and B, because soon we are going to replace B with A.

Recall that for the Bell state $\mid \Psi^- \rangle_{AB}$ qubit B is always an inverted image of qubit A. When measured, it always points in the opposite direction, regardless of which direction we choose.

So we may hypothesize that if we were to measure qubit A first, this would force qubit B automatically into a reverse image of A. Therefore performing a measurement on A first and then on B is really like performing a measurement on A, then flipping what comes out and measuring it again the way that B would be measured.

But we could just as well reverse the order of measurements on A and B or perform both measurements simultaneously and the result in terms of probabilities still ought to be the same, as long as there isn't enough time between one and the other measurement for the other qubit to interact with the environment, since this may change the state of the other qubit in a way that no longer depends on the first qubit only.

Some quantum mechanics purists object to this reasoning saying that quantum mechanics does not allow for a single qubit to be measured simultaneously against polarizers at two different angles. And this is indeed true. But here the same quantum mechanics provides us with a mechanism that lets us overcome this restriction: this mechanism is entanglement, entanglement that makes qubit B an inverted copy of qubit A. So by performing simultanous measurements on A and B we can, in effect, measure, say, A against two different angles at the same time.

Let us then substitute a counter-aligned qubit A in place of qubit B:

$$\nearrow_B \quad \longrightarrow \quad \swarrow_A$$
$$\rightarrow_B \quad \longrightarrow \quad \leftarrow_A$$

But if qubit A passes through the $\swarrow$ polarizer, it is the same as to say that it would *not* pass through the $\nearrow_A$ polarizer, so, in terms of the actual probabilities, we can write:

$$p^{\swarrow_A} = p^{\neg\nearrow_A}$$

where $\neg$ is a Boolean not[4].

Similarly

$$p^{\leftarrow_A} = p^{\neg\rightarrow_A}$$

In effect, our original expression (5.10) can be rewritten as

$$p^{\uparrow_A, \neg\nearrow_A} + p^{\nearrow_A, \neg\rightarrow_A} \tag{5.11}$$

The first term of (5.11) is the probability that qubit A passes through the $\uparrow$ polarizer and at the same time fails to pass through the $\nearrow$ polarizer. The second term is the probability that qubit A passes through the $\nearrow$ polarizer and at the same time fails to pass through the $\rightarrow$ polarizer.

The following, though as it will turn out incorrect, argument asserts that the sum of the two probabilities should be no less than the probability that qubit A passes through the $\uparrow$ polarizer and at the same time fails to pass through the $\rightarrow$ polarizer, $p^{\uparrow_A, \neg\rightarrow_A}$. Why should it be so?

We may conjecture that a qubit that goes through the $\uparrow$ polarizer and at the same time fails to pass through the $\rightarrow$ one, let us call them qubits of the $\{\uparrow, \neg\rightarrow\}$-category[5], *may or may not* pass through the $\nearrow$ polarizer – if we were to subject it to such a measurement. Strictly speaking, in order to implement this measurement we would have to produce a strong entanglement of three qubits, such that they would be always found in identical (or reversed) states. Anyhow, let us get back to our duplicated qubit A. The reasoning goes that if the qubit were to pass through the $\nearrow$ polarizer (this would make it of a $\{\nearrow\}$-category), it would contribute to the $\{\nearrow, \neg\rightarrow\}$-category and thus also to $p^{\nearrow_A, \neg\rightarrow_A}$, and if it were not to pass through the $\nearrow$ polarizer, it would contribute to the $\{\uparrow, \neg\nearrow\}$-category and thus also to $p^{\uparrow_A, \neg\nearrow_A}$.

This is illustrated in figure 5.8.

---

[4]Recall that the fiducial vector of probabilities is always normalized so that for the principal direction, e.g., $\nearrow$ we have that $p^{\nearrow_A} + p^{\swarrow_A} = p^{\nearrow_A} + p^{\neg\nearrow_A} = 1$.

[5]We are using here the word "category" in its traditional rather than mathematical sense, i.e., a distinct class to which entities belong. It is OK to think of these as sets.
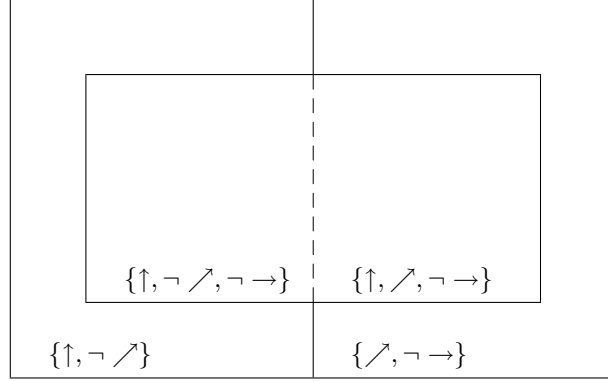
Figure 5.8: A classical look at the Bell inequality. We have three boxes labeled by the categories, with qubits belonging to one or more. Boxes labeled by $\{\uparrow, \neg \nearrow\}$ and $\{\nearrow, \neg \rightarrow\}$ are *disjoint*. The inner box corresponds to qubits in the $\{\uparrow, \neg \rightarrow\}$-category. We subdivide it into two *disjoint* boxes labeled by $\{\uparrow, \neg \nearrow, \neg \rightarrow\}$ and $\{\uparrow, \nearrow, \neg \rightarrow\}$. Qubits in the left box, labeled by $\{\uparrow, \neg \nearrow, \neg \rightarrow\}$ also belong to the $\{\uparrow, \neg \nearrow\}$ box and qubits in the right box, labeled by $\{\uparrow, \nearrow, \neg \rightarrow\}$ also belong to the $\{\nearrow, \neg \rightarrow\}$ box and so we find that the number of qubits N in each category must satisfy: $N^{\uparrow, \neg \nearrow} + N^{\nearrow, \neg \rightarrow} \geq N^{\uparrow, \neg \nearrow, \neg \rightarrow} + N^{\uparrow, \nearrow, \neg \rightarrow} = N^{\uparrow, \neg \rightarrow}$, which translates into $p^{\uparrow_A, \neg \nearrow_A} + p^{\nearrow_A, \neg \rightarrow_A} \geq p^{\uparrow_A, \neg \rightarrow_A}$

In summary:

$$p^{\uparrow_A, \neg \nearrow_A} + p^{\nearrow_A, \neg \rightarrow_A} \geq p^{\uparrow_A, \neg \rightarrow_A}$$

Now, let us substitute qubit B back in place of qubit A:

$$p^{\uparrow_A, \nearrow_B} + p^{\nearrow_A, \rightarrow_B} \geq p^{\uparrow_A, \rightarrow_B} \tag{5.12}$$

But this contradicts equation (5.9) where we have found that

$$p^{\uparrow_A, \nearrow_B} + p^{\nearrow_A, \rightarrow_B} < p^{\uparrow_A, \rightarrow_B}$$

What is amiss?

Inequality (5.12) is one of the celebrated Bell inequalities. Bell and others, e.g., Clauser, Holt, Horne and Shimony [11], and more recently Greenberger, Horne and Zeilinger [24] [23] and Hardy [25] demonstrated several such inequalities and other algebraic expressions that purport to flesh out the difference between classical and quantum physics.

Let us have a closer look at the reasoning that led us to inequality (5.12). The way we transformed biqubit probabilities into single-qubit probabilities was actually quite OK, even though it looked somewhat unnerving. The resulting probabilities all check out. What does not check out is the statement that

> We may conjecture that a qubit that goes through the $\uparrow$ polarizer and at the same time fails to pass through the $\rightarrow$ one ... *may or may not* pass through the $\nearrow$ polarizer .... If the qubit were to pass through the $\nearrow$ polarizer ... it would contribute to the $\{\nearrow, \neg \rightarrow\}$-category and thus also to $p^{\nearrow_A, \neg\rightarrow_A}$, and if it were not to pass through the $\nearrow$ polarizer, it would contribute to the $\{\uparrow, \neg \nearrow\}$-category and thus also to $p^{\uparrow_A, \neg\nearrow_A}$.

What fails in quantum physics is the division of the set of qubits in the $\{\uparrow, \neg \rightarrow\}$-category into two *disjoint* groups depending on whether the qubit *would or would not* pass through the $\nearrow$ polarizer if we were to carry out this additional measurement. This division is *hypothetical* because we do not actually carry out this measurement. But if we do not carry out this measurement then the property in question does not exist and so it cannot be used to split the set.

What does exist is vector $\boldsymbol{r}$ that describes the quantum state of each qubit in the $\{\uparrow, \neg \rightarrow\}$-category. If this vector points in a direction other than $\nearrow$ or $\swarrow$, and we may expect this to be the case for *almost* every qubit in this category, then there is a non-zero probability that the qubit will pass through the $\nearrow$ polarizer, but there is also a non-zero probability that it will not. So *the same qubit* in the $\{\uparrow, \neg \rightarrow\}$-category may contribute to both the $\{\uparrow, \neg \nearrow\}$-category and to the $\{\nearrow, \neg \rightarrow\}$-category, which are disjoint, because they are based on the actual measurement. This has the effect of *swelling the volume* of the $\{\uparrow, \neg \rightarrow\}$-category with respect to this hypothetical $\nearrow$ measurement, and so in some cases we end up with the violation of the Bell inequality.

The Bell inequality (5.12) holds for some angles, in particular it holds for θ_{AB} of $0°$, $90°$, and $180°$, i.e., the angles we have in our basic fiducial matrix of a biqubit[6], but it does not hold for some other angles, most notably for the combination of θ_{AB} of $45°$ and $90°$.

What is it then that the Bell inequality (5.12) and its violation in quantum physics illustrated by inequality (5.9) tell us?

---

[6]This is why the problem was not noticed in 1935 when Einstein, Podolsky and Rosen first considered a biqubit measurement of the kind discussed here. It was only in 1964 that Bell noticed the discrepancy between predictions of quantum mechanics and predictions based on "local realism" and encapsulated the discrepancy in the form of experimentally testable inequalities.

They tell us more than one thing, and this is the problem. The discrepancy is the result of various aspects of quantum physics combining, which leads to confusion.

At first glance we may draw a conclusion that the physical process underlying the projection of a qubit in some quantum state onto a specific direction of a polarizer is truly random, i.e., that there cannot be a hidden switch inside the qubit that *determines* (with 100% accuracy) how the qubit is going to align. If there was such a switch inside, then we could base our division of the $\{\uparrow, \neg \rightarrow\}$-category into the two disjoint $\{\uparrow, \neg \nearrow, \neg \rightarrow\}$ and $\{\uparrow, \nearrow, \neg \rightarrow\}$ categories based on the state of this switch and so the inequality would hold. In quantum mechanics parlance we say that *no deterministic local hidden parameters theory can explain the quantum mechanical result* – the "local hidden parameter" being the switch that sits inside the qubit.

But this result does not exclude the possiblity of non-deterministic local hidden parameters and it does not exclude the possiblity of deterministic non-local hidden parameters either.

We do have a "local hidden parameter" inside a qubit. It is vector $\boldsymbol{r}$. It is "hidden", because we cannot get at it in a single measurement. To evaluate all three components of $\boldsymbol{r}$ we must explore the whole statistical ensemble of the qubit with instruments designed to measure them and no instrument can measure all three at once either. This parameter, however, does not *determine* how a qubit is going to behave when measured. It only provides us with probabilities of various outcomes. So it is not a deterministic parameter.

> We also have a deterministic non-local theory, much favoured by Bell [4], that can be used to reproduce violations of Bell inequalities and other similar expressions [14], which is due primarily to Erwin Schrödinger (1887-1961), Louis de Broglie (1982-1987), David Bohm (1917-1992) and Basil Hiley [8]. This theory has some troubling implications and is not well known. Some physicists may have heard of it, but few studied it in depth. This is a great pity, because the theory is physically and logically unassailable on account of being derived entirely and solely from the fundamental equations of quantum mechanics, so all that quantum mechanics predicts, this theory predicts too.
>
> The theory provides "classical dynamic explanations" for all quantum phenomena such as spin, probabilities, measurements, collapse of the wave function, interference fringes, non-locality, etc., and, who knows, ... it may even be true, sic! But true or not, it makes good reading and gives one plenty of food for thought.
>
> The troubling implication of the theory is that on the fundamental level it treats the whole universe like a single indivisible object: nothing can be truly isolated from the rest. Such isolation and subsequent identification of the isolated components as, e.g., "individual electrons", becomes possible

in thermodynamic limit only, hence the title of Bohm's and Hiley's book "The Undivided Universe". The other troubling implication is the violation of special relativity on the fundamental level, but special relativity is then recovered on the level of expectation values.

Yet, these are the implications of quantum theory taken to its logical conclusion, even if somewhat beyond the point where most physicists are prepared to go – and, needless to say, we have already arrived at quite similar conclusions when studying the inseparability of a biqubit. These will be explored further in the next section.

Why this particular picture should be more troubling to the majority of physicists than, say, multiple universes, for which there is not a shred of experimental evidence, or geometric dimensions in excess of 3+1, for which there is not a shred of experimental evidence either, is hard to tell. If, as Einstein commented, Bohm's theory (not called "Bohm's" back then) is *too cheap*, the other theories seem far *too expensive*. At the same time they all, including Bohm's, suffer from the same fundamental malady of dragging macroscopic, classical concept of space-time into their framework.

5.7   Nonlocality

When applied to a biqubit system in the Bell state $\mid \Psi^- \rangle_{AB}$ the violation of Bell inequality (5.12) illustrates one more aspect of quantum physics: the non-locality of a biqubit.

What does it mean?

The quantum mechanical description of a biqubit is extremely primitive. There is nothing here about the actual physical location of the biqubit components. Suppose, both components are separated by a large *macroscopic* distance, for example 600 m as has been demonstrated in a fairly recent photon experiment by Aspelmeyer and his 12 colleagues from Institut für Experimentalphysik, Universität Wien in Austria [2] – or even a planetary-scale distance as is planned for a forthcoming satellite-based experiment. Can it really be that qubit B measured characteristics end up being *always* opposite of qubit A's – if, as the Bell inequality tells us, they are made quite at random at the point of the measurement rather than due to some "deterministic hidden parameter" inside the qubit? How can qubit B know *instantaneously* what state qubit A has been filtered into? Surely, there ought to be some retardation terms inserted into the probability matrix, to the effect that qubit B would learn about qubit A's encounter with the polarization filter after, say, x_{AB}/c only, where x_{AB} is the distance between the qubits and c is the speed of light.

This is exactly the objection that Einstein, Podolsky and Rosen brought up in 1935 [16]. Their conclusion was that since no information could travel faster than light, there had to be a "deterministic hidden parameter" inside a qubit that would predetermine the way both qubits would interact with the polarizing filters. The Bell inequalities and numerous subsequent experiments [18] demonstrated clearly that there could not be such a "deterministic hidden parameter" inside a qubit. It is clear that on the most fundamental level quantum mechanics of multiple qubits contradicts special theory of relativity. This fairly superficial observation is further confirmed by a more formal proof provided by Hardy [25] and the realization of Hardy's *gedanken* experiment by Irvine, Hodelin, Simon and Bouwmeester [31].

It is well worth having a brief look at how the Bohm's theory explains what happens.

In the Bohm's theory every quantum particle is associated with a field, called *quantum potential*, that stretches all the way to infinity and does not diminish with distance. The field can be derived from the Schrödinger and Dirac equations on fairly standard substitutions similar to what physicists do within the so called WKB approximation. But in the Bohm theory we don't approximate. We calculate things exactly. Taken to this level, fundamental equations of quantum mechanics can be interpreted as equations that describe congruences of trajectories as determined by the quantum potential and various other externally applied fields, e.g., electric, magnetic, gravitational. The interaction of the particle with the quantum potential is instantaneous meaning that whatever the field "touches" and however far away, has an instantaneous effect on the particle.

A measuring apparatus is also a quantum object and so it has a quantum potential field associated with it too.

A biqubit confronted by two widely spaced polarizers is a system of two qubits and two polarizers all joined with the fabric of the shared non-local quantum potential. Whatever happens to qubit A is immediately and instantaneously transmitted to qubit B and vice versa. But not only this. A configuration of both polarizers is also transmitted to both qubits, even before they arrive at their respective points of measurement, and has an effect on how they align.

A heuristic analysis presented in [8] as well as detailed calculations presented by Durt and Pierseaux [14] show how this mutual and instantaneous coupling of all four partners results in the violation of Bell inequalities and other similar expressions.

This model is quite telling. It not only points to the instantaneous interaction between qubits A and B, something that we are forced to expect as soon as we learn about the entanglement and find about its various experimental demonstrations, but, just as importantly, it tells us that we must be careful when thinking about

the act of measurement itself. It is a physical interaction between two or more physical systems. If all participants to the measurement have non-local feelers, we have to consider the possibility of a biqubit adjusting itself to a configuration of the measuring instruments prior to the actual measurement.

The resulting "conspiracy of nature" gives us the quantum reality that puzzles us so at every step.

It is not necessary to *believe* in the Bohm's theory to appreciate various important and interesting points this model makes, just as it is not necessary to be a Christian to appreciate the wisdom of Christ's parables.

How else can we think of an entangled biqubit?

Another approach would be to be more radical about the very notion of space-time itself. After all space-time is a macroscopic construction that requires macroscopic rulers and clocks to define. But we can't take these into the quantum domain, so we should not drag the classical fabric of space-time into the quantum domain either. Yet this is what just about all present-day theories do, perhaps with the notable exception of Smolin's "loop quantum gravity" [49] and "spin networks" of Roger Penrose [43].

How could we replace classical macroscopic space-time in the quantum domain? For example, we could think of a graph of interactions. Quantum systems, e.g., qubits, biqubits, and n-qubits would not be embedded in any space-time. There would not be any distance between them. Instead they would exchange various properties with each other. In some cases the exchanges would be intense, in other cases they would be weak. The exchanges might be ordered too, though not necessarily strongly. In the thermodynamic limit the graph may turn into space and time. The weakly or seldom interacting quantum objects might appear as being far away from each other. The strongly or frequently interacting ones, as being close. The ordering of the graph may turn into macroscopic time. Projections of the interactions between quantum systems onto this macroscopic time and space may acquire some randomness, which is what we, the Creatures of Macroscopia, see when we look at the quantum world.

Everything in this system would stay together at the most fundamental level, and this would explain the non-locality of biqubits – their eventual separation into two independent particles occurring only at the macroscopic level as the result of interaction with great many other nodes of the graph.

Can it really be that all the universe, all physical reality somehow exists at a single point and its macroscopically observable spacious and temporal extent is an illusion built from myriads of interactions within this point?

Why not? Consider photons. In the photon's system of reference time stops because of relativistic time dilation. A photon does not experience time. Similarly, in the photon's system of reference the whole world is squeezed into a point because of Lorentz contraction. From the photon's personal point of view it is everywhere at the same time. Could it be that the photon is right?

In some sense this picture is not so distant from Bohm's theory. It is like Bohm's theory with the space and time taken out of it, so that its quantum potential *can* be non-local.

It is not uncommon in theoretical physics that different conceptual and mathematical frameworks turn out to be equivalent and lead to identical physics.

At the end of the day none of the above may be true. But it is certainly true that Bell inequalities and quantum physics force us to radically revise our often naive notions about the nature of reality.

5.8   Single qubit expectation values

Consider a separable biqubit state described by $\boldsymbol{p}_A \otimes \boldsymbol{p}_B$. It is easy to extract just one of the probabilities from it. We did something similar when defining energy form for a system of two separate non-interacting qubits in section 5.1, page 181.

The trick is to contract the biqubit with $\boldsymbol{\varsigma}^1_{A,B}$ and make use of $\langle \boldsymbol{\varsigma}^i, \boldsymbol{\varsigma}_j \rangle = 2\delta^i{}_j$, where $i, j = 1, x, y, z$. And so:

$$\langle \boldsymbol{\varsigma}^1_B, \boldsymbol{p}_A \otimes \boldsymbol{p}_B \rangle = \boldsymbol{p}_A$$
$$\langle \boldsymbol{\varsigma}^1_A, \boldsymbol{p}_A \otimes \boldsymbol{p}_B \rangle = \boldsymbol{p}_B$$

We can think of $\langle \boldsymbol{\varsigma}^1_B, \boldsymbol{p}_A \otimes \boldsymbol{p}_B \rangle$ as an expectation value for the measurement $\boldsymbol{\varsigma}^1_B$ on $\boldsymbol{p}_A \otimes \boldsymbol{p}_B$. Translating this into the language of quaternions, we find:

$$\boldsymbol{\rho}_A = \boldsymbol{\rho}_A 2\Re \left(\mathbf{1}_B \cdot \boldsymbol{\rho}_B \right) = \boldsymbol{\rho}_A 2\Re \boldsymbol{\rho}_B$$

and now, switching from quaternions to Pauli matrices,

$$\boldsymbol{\rho}_A = \boldsymbol{\rho}_A \mathrm{Tr}_B \boldsymbol{\rho}_B$$

This last expression can be rewritten as

$$\boldsymbol{\rho}_A = \mathrm{Tr}_B \boldsymbol{\rho}_{AB}$$

where for a separable biqubit $\boldsymbol{\rho}_{AB} = \boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B$. The symbol Tr_B means "taking trace over variables that pertain to qubit B" and is referred to as a *partial trace* operation. Physicists also talk about *tracing particle B out* of the biqubit state.

It is interesting to *trace qubit B out* of the Bell state $\mid \Psi^- \rangle_{AB}$. Recall that

$$\mid \Psi^- \rangle_{AB} \equiv \frac{1}{4} \left(\varsigma_{1A} \otimes \varsigma_{1B} - \varsigma_{xA} \otimes \varsigma_{xB} - \varsigma_{yA} \otimes \varsigma_{yB} - \varsigma_{zA} \otimes \varsigma_{zB} \right)$$

Contracting it with ς_B^1 yields

$$\frac{1}{4} \left(\varsigma_{1A} \cdot 2 \right) = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

which implies that $r_A = 0$. This is a completely depolarized state, a total mixture. Similarly, we'd find that particle B, when looked at separately from particle A appears completely depolarized. Yet the biqubit itself is thoroughly polarized. It is in a *pure* state.

There are several interesting conclusions that can be drawn from this observation.

The first conclusion is that this particular biqubit state, the Bell state $\mid \Psi^- \rangle_{AB}$, does not provide us with means of transferring useful information from a point where qubit A is measured to a point where qubit B is measured. At every point of measurement the measured qubit, be it A or B, appears completely chaotic. It is only afterwards, when the results of the measurements for qubits A and B are *compared* that we get to realize that the two qubits were entangled and in a *pure* biqubit state.

The second conclusion is that if we were to associate specific information with the biqubit, as we associated 0 with $|\uparrow\rangle$ and 1 with $|\downarrow\rangle$ previously, the information would be contained in the *entanglement*, i.e., in the biqubit correlations and not in the individual qubits of the system.

The third conclusion is that although the unitary formalism does not, at first glance, let us discuss mixtures and therefore the act of measurement either, here we have a unitary model that captures a mixed state too. It does it by entangling a qubit with another qubit. The result of the entanglement is that the state of each individual qubit becomes mixed. We can therefore generate mixed states of quantum subsystems within the unitary formalism by viewing them as part of larger unitary systems. And we can try to model the act of measurement by entangling the measured quantum object with the measuring apparatus – the quantum state of the whole being unitary, but the quantum state of the measured object decaying into a mixture.

We will discuss this in more detail later, but first let us apply the $\varsigma_{A,B}^1$ measurement to the most general biqubit state, which is, as we saw earlier, is

$$
\boldsymbol{p}_{AB} \;=\; \frac{1}{4}\left(\varsigma_{1A} \otimes \varsigma_{1B} + \sum_{i=x,y,z} r_A^i\, \varsigma_{iA} \otimes \varsigma_{1B} + \sum_{i=x,y,z} r_B^i\, \varsigma_{1A} \otimes \varsigma_{iB} \right.
$$
$$
\left. + \sum_{i,j=x,y,z} x_{AB}^{ij}\, \varsigma_{iA} \otimes \varsigma_{jB} \right) \tag{5.13}
$$

Contracting this with ς_B^1 yields

$$
\boldsymbol{p}_A = \langle \varsigma_B^1, \boldsymbol{p}_{AB} \rangle = \frac{1}{2}\left(\varsigma_{1A} + \sum_{i=x,y,z} r_A^i \varsigma_{iA} \right)
$$

and contracting $\boldsymbol{p}_{AB}$ with ς_A^1 yields

$$
\boldsymbol{p}_B = \langle \varsigma_A^1, \boldsymbol{p}_{AB} \rangle = \frac{1}{2}\left(\varsigma_{1B} + \sum_{i=x,y,z} r_B^i \varsigma_{iB} \right)
$$

These two formulas give us a new interpretation of $r_{A,B}^i$ in a biqubit. The biqubit coefficients $r_{A,B}^i$ encode the results of *separate* measurements on qubits A and B. We can read these directly and easily from equation (5.13). Because they correspond to individual qubits at both ends of a biqubit, they must both lie within the respective single qubit Bloch balls, i.e., we must have that $\boldsymbol{r}_A \cdot \boldsymbol{r}_A \leq 1$ and $\boldsymbol{r}_B \cdot \boldsymbol{r}_B \leq 1$. We have seen previously that these coefficients mix the same way they do in single qubit systems.

This again confirms that if we focus on one component of a biqubit and ignore the other one, we have no way of telling whether this qubit is entangled with another qubit or just mixed for some other reason.

5.9   Classification of biqubit states

A biqubit system appears very simple at first glance. Yet, when investigated in more depth, it reveals a great deal of complexity. The complexity derives primarily from the many ways in which biqubits can be mixed: both on the level of individual qubits, which is described by $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$, and on the level of biqubits themselves, which is captured by $\boldsymbol{x}_{AB}$. On top of this we have pure states and entangled states, and the latter can be mixed too.

Yet, it is still so surprising to learn that most work on biqubit separability and classification began in the mid-1990s only, some 70 years after the birth of quantum mechanics and after Pauli's discovery of his exclusion principle. The reason for this is that prior to that time, most physicists considered biqubits within the confines of unitary formalism, and here matters are greatly simplified. It was the new interest in Bell inequalities and related biqubit probability measurements that made physicists ask whether the correlations they observed were indeed of quantum nature or merely due to classical mixing. This very fundamental question proved remarkably difficult to answer.

At the same time, this topic is central to quantum computing, because everything in quantum computing is done with biqubits. The reason for this is that a biqubit gate, called the *controlled-not* gate is universal to quantum computing. And the other reason is that the moment we enter the domain of experimental quantum physics we have to abandon the comfortable, but idealistic world of the unitary formalism and face the reality of fully blown probability theory, or the *density operator* theory as the physicists prefer to call it in this context. As working cars cannot be designed without taking friction into account, similarly, working quantum computers cannot be designed without taking depolarization, dissipation and other non-unitary phenomena into account. We saw it very clearly in the quantronium example.

Let us go back to the biqubit representation given by equation (5.7):

$$
\boldsymbol{p}_{AB} = \frac{1}{4}\left(\varsigma_{1A} \otimes \varsigma_{1B} + \sum_{j=x,y,z} r_B^j \varsigma_{1A} \otimes \varsigma_{jB} + \sum_{i=x,y,z} r_A^i \varsigma_{iA} \otimes \varsigma_{1B} + \sum_{i,j=x,y,z} x_{AB}^{ij} \varsigma_{iA} \otimes \varsigma_{jB}\right) \tag{5.14}
$$

where we have normalized the probabilities so that α, the coefficient in front of $\varsigma_{1A} \otimes \varsigma_{1B}$, is 1.

We have seen in section 5.5 how to change the frame in which a biqubit is measured and what effect this has on equation (5.14). The basic idea there was that the three Pauli vectors labeled by x, y and z behaved under rotations like normal 3-dimensional vectors that pointed in the x, y and z directions.

Equation (5.14) evaluates probabilities $\boldsymbol{p}_{AB}$ in terms of Pauli vectors attached at two different locations, the location of qubit A and the location of qubit B. These two Pauli frames don't have to be oriented the same way and can be rotated independently of each other – exactly as we did in section 5.5.

An arbitrary rotation in the 3D space can be characterized by providing three Euler angles. The two independent rotations, one for qubit A and the other one for qubit B, are therefore specified by six Euler angles. We can always choose the six Euler angles so as to kill the six off-diagonal elements of matrix x_{AB}^{ij}.

> Matrix x_{AB}^{ij} does not have to be symmetric for this and we will not end up with complex numbers on the diagonal either, because here we manipulate both frames independently. If we wanted to diagonalize matrix x_{AB}^{ij} by performing an *identical* rotation on both frames, then the matrix would have to be symmetric for this to work. This is because then we would have only *three* Euler angles to play with, and with these we could only kill *three* off-diagonal elements.

Having diagonalized matrix x_{AB}^{ij} we end up with only 9 real numbers (that aren't zero) in $\boldsymbol{p}_{AB}$ in place of the original 15. And so, it turns out that of the 15 degrees of freedom that characterize the biqubit in the fiducial formalism, 6 are of purely geometric character and can be eliminated or otherwise modified by rotating frames against which the biqubit components are measured. But the remaining 9 degrees of freedom are physical.

Once we have x_{AB}^{ij} in the diagonal form, we can switch around the labels on the directions x, y and z so as to rewrite the $\sum_{ij} x_{AB}^{ij} \varsigma_{iA} \otimes \varsigma_{jB}$ term in the following form:

$$\text{sign} \left(\det \boldsymbol{x}_{AB} \right) \left(\kappa_{AB}^x \varsigma_{xA} \otimes \varsigma_{xB} + \kappa_{AB}^y \varsigma_{yA} \otimes \varsigma_{yB} + \kappa_{AB}^z \varsigma_{zA} \otimes \varsigma_{zB} \right)$$

where the *kappa* coefficients are ordered as follows:

$$\kappa_{AB}^x \geq \kappa_{AB}^y \geq \kappa_{AB}^z$$

The κ coefficients are the same as the x_{AB}^{ii} coefficients after the diagonalization up to a sign and ordering.

But even now we may have some freedom left. For example, if all κ_{AB}^i are zero, then we can rotate both frames as much as we wish without changing x_{AB}^{ij} or κ_{AB}^i at all. Furthermore, we're still left with the freedom to reflect rather than rotate the varsigmas. For example:

$$\varsigma_{xA} \rightarrow -\varsigma_{xA} \quad \text{and}$$
$$\varsigma_{xB} \rightarrow -\varsigma_{xB}$$

leaves κ_{AB}^x unchanged.

When such additional freedoms are left after diagonalization of x_{AB}^{ij} we use them to kill as many remaining $r_{A,B}^i$ as possible, usually starting with $r_{A,B}^y$, then proceeding to $r_{A,B}^z$.

The purpose of all these manipulations is to "normalize" the qubit's representation, and remove any dependence on geometry and choices of directions.

The resulting probability matrix is:

$$\boldsymbol{p}_{AB} = \frac{1}{4}\left(\varsigma_{1A} \otimes \varsigma_{1B} + \sum_{i=x,y,z} r_A^i \varsigma_{iA} \otimes \varsigma_{1B} + \sum_{j=x,y,z} r_B^j \varsigma_{1A} \otimes \varsigma_{jB} \right.$$
$$\left. + \operatorname{sign}\left(\det \boldsymbol{x}_{AB}\right) \sum_{k=x,y,z} \kappa_{AB}^k \varsigma_{kA} \otimes \varsigma_{kB} \right) \tag{5.15}$$

It should be remembered that coefficients r_A^i and r_B^j are no longer the same they were in the original version of $\boldsymbol{p}_{AB}$. Rotating Pauli vectors $\varsigma_{iA,B}$ changes not only x_{AB}^{ij}, but r_A^i and r_B^j too.

Now, following Englert and Metwally [17], we can divide all possible biqubit states into classes:

Class A defined by $\kappa_{AB}^x = \kappa_{AB}^y = \kappa_{AB}^z = 0$. The completely chaotic state with all coefficients equal zero belongs to this class. Since we end up with $x_{AB}^{ij} = 0$ for this class we can still rotate both frames, at A and B, and kill $r_{A,B}^{y,z}$, which leaves:

$$\boldsymbol{p}_{AB} = \frac{1}{4}\left(\varsigma_{1A} \otimes \varsigma_{1B} + r_A^x \varsigma_{xA} \otimes \varsigma_{1B} + r_B^x \varsigma_{1A} \otimes \varsigma_{xB} \right)$$

$$\begin{pmatrix} p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\ p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\ p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\ p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes} \end{pmatrix}$$

$$= \frac{1}{4}\begin{pmatrix} 1 & 1 & 1+r_B^x & 1 \\ 1 & 1 & 1+r_B^x & 1 \\ 1+r_A^x & 1+r_A^x & 1+r_A^x+r_B^x & 1+r_A^x \\ 1 & 1 & 1+r_B^x & 1 \end{pmatrix}$$

where $r_A^x \geq 0$ and $r_B^x \geq 0$ too. This is the most general normalized state in this class. It is a very strange state with well defined local single qubit states, but with no biqubit correlations at all, not even classical ones of the form $r_A^x r_B^x \varsigma_{xA} \otimes \varsigma_{xB}$.

Class B$^+$ defined by $\kappa_{AB}^x = \kappa_{AB}^y = \kappa_{AB}^z = \kappa > 0$ and $\det \boldsymbol{x}_{AB} > 0$.

Class B$^-$ defined by $\kappa^x_{AB} = \kappa^y_{AB} = \kappa^z_{AB} = \kappa > 0$ and $\det \boldsymbol{x}_{AB} < 0$.

Here we still have the freedom to kill additionally three $r^i_{A,B}$. Following an established convention we choose

$$
\begin{aligned}
\boldsymbol{p}_{AB} \;=\; \frac{1}{4}\Big(&\varsigma_{1A} \otimes \varsigma_{1B} \\
&+ r^x_A \varsigma_{xA} \otimes \varsigma_{1B} \\
&+ \varsigma_{1A} \otimes (r^x_B \varsigma_{xB} + r^z_B \varsigma_{zB}) \\
&\pm \kappa \left(\varsigma_{xA} \otimes \varsigma_{xB} + \varsigma_{yA} \otimes \varsigma_{yB} + \varsigma_{zA} \otimes \varsigma_{zB} \right) \Big)
\end{aligned}
$$

$$
\begin{pmatrix}
p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\
p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\
p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\
p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes}
\end{pmatrix}
$$
$$
= \frac{1}{4}
\begin{pmatrix}
1 + r^z_B \pm \kappa & 1 - r^z_B \mp \kappa & 1 + r^x_B & 1 \\
1 + r^z_B \mp \kappa & 1 - r^z_B \pm \kappa & 1 + r^x_B & 1 \\
1 + r^x_A + r^z_B & 1 + r^x_A - r^z_B & 1 + r^x_A + r^x_B \pm \kappa & 1 + r^x_A \\
1 + r^z_B & 1 - r^z_B & 1 + r^x_B & 1 \pm \kappa
\end{pmatrix}
$$

where $r^x_A \geq 0$ and when $r^x_A > 0$ then $r^z_B \geq 0$ and when $r^x_A = 0$ then $r^x_B \geq 0$ and $r^z_B = 0$.

Two important families of states:

$$
\frac{1}{4}\left(\varsigma_{1A} \otimes \varsigma_{1B} \pm \kappa \left(\varsigma_{xA} \otimes \varsigma_{xB} + \varsigma_{yA} \otimes \varsigma_{yB} + \varsigma_{zA} \otimes \varsigma_{zB} \right) \right)
$$

belong to class B$^\pm$. They are called *Werner states*. A B$^-$ Werner state with $\kappa = 1$ is the Bell state $\mid \Psi^- \rangle_{AB}$.

Class C defined by $\kappa^x_{AB} = \kappa > \kappa^y_{AB} = \kappa^z_{AB} = 0$.

Here $\det \boldsymbol{x}_{AB} = 0$, because $\kappa^y_{AB} = \kappa^z_{AB} = 0$. This leaves us with enough freedom to clean up $r^y_{A,B}$ too – such is the established choice – so that the state looks as follows:

$$
\begin{aligned}
\boldsymbol{p}_{AB} \;=\; \frac{1}{4}\Big(&\varsigma_{1A} \otimes \varsigma_{1B} \\
&+ (r^x_A \varsigma_{xA} + r^z_A \varsigma_{zA}) \otimes \varsigma_{1B} \\
&+ \varsigma_{1A} \otimes (r^x_B \varsigma_{xB} + r^z_B \varsigma_{zB}) \\
&\pm \kappa \varsigma_{xA} \otimes \varsigma_{xB} \Big)
\end{aligned}
$$

$$
\begin{pmatrix}
p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\
p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\
p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\
p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes}
\end{pmatrix}
$$
$$
= \frac{1}{4}
\begin{pmatrix}
1 + r_A^z + r_B^z & 1 + r_A^z - r_B^z & 1 + r_B^x + r_A^z & 1 + r_A^z \\
1 - r_A^z + r_B^z & 1 - r_A^z - r_B^z & 1 + r_B^x - r_A^z & 1 - r_A^z \\
1 + r_A^x + r_B^z & 1 + r_A^x - r_B^z & 1 + r_A^x + r_B^x \pm \kappa & 1 + r_A^x \\
1 + r_B^z & 1 - r_B^z & 1 + r_B^x & 1
\end{pmatrix}
$$

where $r_{A,B}^z \geq 0$ and $r_A^x \geq 0$ and when $r_A^x = 0$ then $r_B^x \geq 0$.

A simple separable biqubit state $\boldsymbol{p}_A \otimes \boldsymbol{p}_B$ defined solely by $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$ belongs to this class. We can rotate both frames so that $\boldsymbol{r}_A = r_A^x \boldsymbol{e}_{xA}$ and $\boldsymbol{r}_B = r_B^x \boldsymbol{e}_{xB}$. Then

$$
\boldsymbol{p}_{AB} = \frac{1}{4} \left(\varsigma_{1A} \otimes \varsigma_{1B} + r_A^x \varsigma_{xA} \otimes \varsigma_{1B} + r_B^x \varsigma_{1A} \otimes \varsigma_{xB} + r_A^x r_B^x \varsigma_{xA} \otimes \varsigma_{xB} \right)
$$

Class $\mathbf{D}^+$ defined by $\kappa_{AB}^x > \kappa_{AB}^y = \kappa_{AB}^z = \kappa > 0$ and $\det \boldsymbol{x}_{AB} > 0$.

Class $\mathbf{D}^-$ defined by $\kappa_{AB}^x > \kappa_{AB}^y = \kappa_{AB}^z = \kappa > 0$ and $\det \boldsymbol{x}_{AB} < 0$.

Here we have less freedom left after the diagonalization of x_{AB}^{ij} and the only $r_{A,B}^i$ that we can to get rid of is r_A^y. The resulting state looks as follows:

$$
\begin{aligned}
\boldsymbol{p}_{AB} = \quad & \frac{1}{4} \Bigg(\varsigma_{1A} \otimes \varsigma_{1B} \\
& + \left(r_A^x \varsigma_{xA} + r_A^z \varsigma_{zA} \right) \otimes \varsigma_{1B} \\
& + \varsigma_{1A} \otimes \left(r_B^x \varsigma_{xB} + r_B^y \varsigma_{yB} + r_B^z \varsigma_{zB} \right) \\
& \pm \left(\kappa^x \varsigma_{xA} \otimes \varsigma_{xB} + \kappa \left(\varsigma_{yA} \otimes \varsigma_{yB} + \varsigma_{zA} \otimes \varsigma_{zB} \right) \right) \Bigg)
\end{aligned}
$$

$$
\begin{pmatrix}
p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\
p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\
p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\
p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes}
\end{pmatrix}
$$
$$
= \frac{1}{4}
\begin{pmatrix}
1 + r_A^z + r_B^z \pm \kappa & 1 + r_A^z - r_B^z \mp \kappa & 1 + r_A^z + r_B^x & 1 + r_A^z + r_B^y \\
1 - r_A^z + r_B^z \mp \kappa & 1 - r_A^z - r_B^z \pm \kappa & 1 - r_A^z + r_B^x & 1 - r_A^z + r_B^y \\
1 + r_A^x + r_B^z & 1 + r_A^x - r_B^z & 1 + r_A^x + r_B^x \pm \kappa^x & 1 + r_A^x + r_B^y \\
1 + r_B^z & 1 - r_B^z & 1 + r_B^x & 1 + r_B^y \pm \kappa
\end{pmatrix}
$$

where $r_A^x \geq 0$, $r_A^z \geq 0$ and $r_B^y \geq 0$. When $r_A^x = 0$ then $r_B^x \geq 0$. When $r_A^z = 0$ and $r_B^y = 0$ then $r_B^z \geq 0$.

All pure states belong to class D^-. Their generic form is

$$
\boldsymbol{p}_{AB} = \frac{1}{4}\Bigg(\boldsymbol{\varsigma}_{1A} \otimes \boldsymbol{\varsigma}_{1B}
$$
$$
+ r\left(\boldsymbol{\varsigma}_{xA} \otimes \boldsymbol{\varsigma}_{1B} - \boldsymbol{\varsigma}_{1A} \otimes \boldsymbol{\varsigma}_{xB}\right)
$$
$$
- \left(\boldsymbol{\varsigma}_{xA} \otimes \boldsymbol{\varsigma}_{xB} + \sqrt{1-r^2}\left(\boldsymbol{\varsigma}_{yA} \otimes \boldsymbol{\varsigma}_{yB} + \boldsymbol{\varsigma}_{zA} \otimes \boldsymbol{\varsigma}_{zB}\right)\right)\Bigg)
$$

$$(5.16)$$

where $r \in [0, 1]$.

It is easy to see why this must be the generic form of a pure state. Pure states are described by 4 complex numbers, or 8 real numbers, constrained by one normalization condition. This leaves 7 real numbers. But 6 of these can be eliminated by frame rotations, so that we end up with just one generic parameter. This generic parameter is r in the above equation. So the dimensionality is just right.

The next thing that we have to check is that this is indeed a pure state. We can do this by demonstrating that its corresponding density quaternion is idempotent, i.e., that $\boldsymbol{\rho}\boldsymbol{\rho} = \boldsymbol{\rho}$.

Let us begin by replacing varsigmas with sigmas:

$$
\boldsymbol{\rho}_{AB} = \frac{1}{4}\Bigg(\mathbf{1}_A \otimes \mathbf{1}_B
$$
$$
+ r\left(\boldsymbol{\sigma}_{xA} \otimes \mathbf{1}_B - \mathbf{1}_A \otimes \boldsymbol{\sigma}_{xB}\right)
$$
$$
- \boldsymbol{\sigma}_{xA} \otimes \boldsymbol{\sigma}_{xB} - \sqrt{1-r^2}\left(\boldsymbol{\sigma}_{yA} \otimes \boldsymbol{\sigma}_{yB} + \boldsymbol{\sigma}_{zA} \otimes \boldsymbol{\sigma}_{zB}\right)\Bigg)
$$

$$(5.17)$$

Let us organize the computation by introducing

$$
a = \left(\boldsymbol{\sigma}_{xA} \otimes \mathbf{1}_B - \mathbf{1}_A \otimes \boldsymbol{\sigma}_{xB}\right)
$$
$$
b = \boldsymbol{\sigma}_{xA} \otimes \boldsymbol{\sigma}_{xB}
$$
$$
c = \left(\boldsymbol{\sigma}_{yA} \otimes \boldsymbol{\sigma}_{yB} + \boldsymbol{\sigma}_{zA} \otimes \boldsymbol{\sigma}_{zB}\right)
$$

then

$$
\boldsymbol{\rho}_{AB} \cdot \boldsymbol{\rho}_{AB} = \frac{1}{16}\left(\mathbf{1}_A \otimes \mathbf{1}_B + ra - b - \sqrt{1-r^2}c\right)
$$

$$\times \left(\mathbf{1}_A \otimes \mathbf{1}_B + ra - b - \sqrt{1 - r^2}c \right)$$
$$= \frac{1}{16} \Big(\mathbf{1}_A \otimes \mathbf{1}_B + 2ra - 2b - 2\sqrt{1 - r^2}c$$
$$+ r^2 a^2 + b^2 + (1 - r^2)c^2$$
$$- r(ab + ba)$$
$$- r\sqrt{1 - r^2}(ac + ca)$$
$$+ \sqrt{1 - r^2}(bc + cb) \Big)$$

The trick is now to remember that A sigmas must multiply other A sigmas only and the same holds for B sigmas.

It is very easy to see that

$$b^2 = \boldsymbol{\sigma}_{xA} \otimes \boldsymbol{\sigma}_{xB} \cdot \boldsymbol{\sigma}_{xA} \otimes \boldsymbol{\sigma}_{xB} = \mathbf{1}_A \otimes \mathbf{1}_B$$

This is because $\boldsymbol{\sigma}_{iA,B} \boldsymbol{\sigma}_{iA,B} = \mathbf{1}_{A,B}$, for $i = x, y, z$.

It is almost as easy to see that

$$c^2 = 2\left(\mathbf{1}_A \otimes \mathbf{1}_B - \boldsymbol{\sigma}_{xA} \otimes \boldsymbol{\sigma}_{xB} \right) = 2\left(\mathbf{1}_A \otimes \mathbf{1}_B - b \right)$$

This is, first, because of the above, and second, because

$$\boldsymbol{\sigma}_{yA} \otimes \boldsymbol{\sigma}_{yB} \cdot \boldsymbol{\sigma}_{zA} \otimes \boldsymbol{\sigma}_{zB} = \boldsymbol{\sigma}_{zA} \otimes \boldsymbol{\sigma}_{zB} \cdot \boldsymbol{\sigma}_{yA} \otimes \boldsymbol{\sigma}_{yB} = -\boldsymbol{\sigma}_{xA} \otimes \boldsymbol{\sigma}_{xB}$$

The minus here comes from i^2.

And it is childishly easy to see that

$$a^2 = c^2 = 2\left(\mathbf{1}_A \otimes \mathbf{1}_B - b \right)$$

too. This is because

$$\mathbf{1}_A \otimes \boldsymbol{\sigma}_{xB} \cdot \mathbf{1}_A \otimes \boldsymbol{\sigma}_{xB} = \mathbf{1}_A \otimes \mathbf{1}_B$$

and because

$$\mathbf{1}_A \otimes \boldsymbol{\sigma}_{xB} \cdot \boldsymbol{\sigma}_{xA} \otimes \mathbf{1}_B = \boldsymbol{\sigma}_{xA} \otimes \boldsymbol{\sigma}_{xB} = b$$

Let us then add

$$r^2 a^2 + b^2 + (1 - r^2)c^2$$
$$= r^2 2\left(\mathbf{1}_A \otimes \mathbf{1}_B - b \right) + \mathbf{1}_A \otimes \mathbf{1}_B + (1 - r^2)2\left(\mathbf{1}_A \otimes \mathbf{1}_B - b \right)$$
$$= 3\mathbf{1}_A \otimes \mathbf{1}_B - 2b$$

and so

$$\mathbf{1}_A \otimes \mathbf{1}_B + 2ra - 2b - 2\sqrt{1 - r^2}c + r^2 a^2 + b^2 + (1 - r^2)c^2$$
$$= 4\mathbf{1}_A \otimes \mathbf{1}_B + 2ra - 4b - 2\sqrt{1 - r^2}c$$

For perfect happiness we still have to generate additional $2ra$ and additional $-2\sqrt{1-r^2}c$ using the remaining three anti-commutator terms.

It is quite trivial to see that

$$bc = cb = -c$$

This is because

$$\boldsymbol{\sigma}_x \boldsymbol{\sigma}_y = -\boldsymbol{\sigma}_y \boldsymbol{\sigma}_x = i\boldsymbol{\sigma}_z$$

and

$$\boldsymbol{\sigma}_x \boldsymbol{\sigma}_z = -\boldsymbol{\sigma}_z \boldsymbol{\sigma}_x = -i\boldsymbol{\sigma}_y$$

So bc is merely going to swap $\boldsymbol{\sigma}_{y\,A} \otimes \boldsymbol{\sigma}_{y\,B}$ and $\boldsymbol{\sigma}_{z\,A} \otimes \boldsymbol{\sigma}_{z\,B}$ and throw $i^2 = (-i)^2 = -1$ in front. Consequently

$$bc + cb = -2c$$

Similarly

$$ab = ba = -a$$

This is because

$$\boldsymbol{\sigma}_{x\,A} \otimes \mathbf{1}_B \cdot \boldsymbol{\sigma}_{x\,A} \otimes \boldsymbol{\sigma}_{x\,B} = \mathbf{1}_A \otimes \boldsymbol{\sigma}_{x\,B}$$

Consequently

$$ab + ba = -2a$$

and

$$-r(ab + ba) + \sqrt{1-r^2}(bc + cb) = 2ra - 2\sqrt{1-r^2}c$$

Adding this to

$$4\mathbf{1}_A \otimes \mathbf{1}_B + 2ra - 4b - 2\sqrt{1-r^2}c$$

yields

$$4\mathbf{1}_A \otimes \mathbf{1}_B + 4ra - 4b - 4\sqrt{1-r^2}c$$

which when divided by 16 returns the original $\boldsymbol{\rho}_{AB}$.

We are left with one more term, namely,

$$r\sqrt{1-r^2}(ac + ca)$$

and this term vanishes, because here we have just one $\boldsymbol{\sigma}_{x\,A,B}$ from a multiplying one of the $\boldsymbol{\sigma}_{y,z\,A,B}$ from c, first from the left, in ac, and then from the right, in ca. But different sigmas anti-commute, so this kills the whole term.

This computation, although somewhat tedious, is also instructive. Apart from demonstrating that the quaternion of state (5.17) is idempotent and so the state itself is pure, the example also shows how to divide a lengthy computation of this nature into smaller, manageable chunks, and how to perform the computation itself by using quaternion rules only and not Pauli matrices.

Class E$^+$ defined by $\kappa = \kappa_{AB}^x = \kappa_{AB}^y > \kappa_{AB}^z$ and $\det \boldsymbol{x}_{AB} > 0$.

Class E$^-$ defined by $\kappa = \kappa_{AB}^x = \kappa_{AB}^y > \kappa_{AB}^z$ and $\det \boldsymbol{x}_{AB} < 0$.

These two classes are very similar to D$^\pm$. The difference is that whereas previously we had $\kappa = \kappa^y = \kappa^z$, here we have that $\kappa = \kappa^x = \kappa^y$ instead and κ^z is different.

This, as before, lets us kill r_A^y only and so we end up with

$$
\begin{aligned}
\boldsymbol{p}_{AB} \;=\; \frac{1}{4}\bigg(&\varsigma_{1A} \otimes \varsigma_{1B} \\
&+ \left(r_A^x \varsigma_{xA} + r_A^z \varsigma_{zA} \right) \otimes \varsigma_{1B} \\
&+ \varsigma_{1A} \otimes \left(r_B^x \varsigma_{xB} + r_B^y \varsigma_{yB} + r_B^z \varsigma_{zB} \right) \\
&\pm \left(\kappa \left(\varsigma_{xA} \otimes \varsigma_{xB} + \varsigma_{yA} \otimes \varsigma_{yB} \right) + \kappa^z \varsigma_{zA} \otimes \varsigma_{zB} \right) \bigg)
\end{aligned}
$$

$$
\begin{pmatrix}
p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\
p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\
p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\
p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes}
\end{pmatrix}
$$

$$
= \frac{1}{4}
\begin{pmatrix}
1+r_A^z+r_B^z\pm\kappa^z & 1+r_A^z-r_B^z\mp\kappa^z & 1+r_A^z+r_B^x & 1+r_A^z+r_B^y \\
1-r_A^z+r_B^z\mp\kappa^z & 1-r_A^z-r_B^z\pm\kappa^z & 1-r_A^z+r_B^x & 1-r_A^z+r_B^y \\
1+r_A^x+r_B^z & 1+r_A^x-r_B^z & 1+r_A^x+r_B^x\pm\kappa & 1+r_A^x+r_B^y \\
1+r_B^z & 1-r_B^z & 1+r_B^x & 1+r_B^y\pm\kappa
\end{pmatrix}
$$

where $r_A^x \geq 0$ and $r_B^y \geq 0$ and when $r_A^x = 0$ then $r_B^y = 0$ too and $r_B^x \geq 0$. Additionally $r_A^z \geq 0$ too and when $r_A^z = 0$ then $r_B^z \geq 0$.

Class F$^+$ defined by $\kappa_{AB}^x > \kappa_{AB}^y > \kappa_{AB}^z$ and $\det \boldsymbol{x}_{AB} > 0$.

Class F$^-$ defined by $\kappa_{AB}^x > \kappa_{AB}^y > \kappa_{AB}^z$ and $\det \boldsymbol{x}_{AB} < 0$.

With all three kappas different we get no freedom to kill any r-s. We can deploy reflections in order to make as many of r_A^x, r_B^x, r_A^y, r_B^y, r_A^z and r_B^z as possible positive – in preference of the order listed.

The resulting probability matrix is

$$
\begin{aligned}
\boldsymbol{p}_{AB} \;=\; \frac{1}{4}\bigg(&\varsigma_{1A} \otimes \varsigma_{1B} \\
&+ \left(r_A^x \varsigma_{xA} + r_A^y \varsigma_{yA} + r_A^z \varsigma_{zA} \right) \otimes \varsigma_{1B}
\end{aligned}
$$

$$+\varsigma_{1A} \otimes \left(r_B^x \varsigma_{xB} + r_B^y \varsigma_{yB} + r_B^z \varsigma_{zB} \right)$$

$$\pm \left(\kappa^x \varsigma_{xA} \otimes \varsigma_{xB} + \kappa^y \varsigma_{yA} \otimes \varsigma_{yB} + \kappa^z \varsigma_{zA} \otimes \varsigma_{zB} \right) \Big)$$

$$\begin{pmatrix} p^{\uparrow\uparrow} & p^{\uparrow\downarrow} & p^{\uparrow\rightarrow} & p^{\uparrow\otimes} \\ p^{\downarrow\uparrow} & p^{\downarrow\downarrow} & p^{\downarrow\rightarrow} & p^{\downarrow\otimes} \\ p^{\rightarrow\uparrow} & p^{\rightarrow\downarrow} & p^{\rightarrow\rightarrow} & p^{\rightarrow\otimes} \\ p^{\otimes\uparrow} & p^{\otimes\downarrow} & p^{\otimes\rightarrow} & p^{\otimes\otimes} \end{pmatrix}$$

$$= \frac{1}{4} \begin{pmatrix} 1 + r_A^z + r_B^z \pm \kappa^z & 1 + r_A^z - r_B^z \mp \kappa^z & 1 + r_A^z + r_B^x & 1 + r_A^z + r_B^y \\ 1 - r_A^z + r_B^z \mp \kappa^z & 1 - r_A^z - r_B^z \pm \kappa^z & 1 - r_A^z + r_B^x & 1 - r_A^z + r_B^y \\ 1 + r_A^x + r_B^z & 1 + r_A^x - r_B^z & 1 + r_A^x + r_B^x \pm \kappa^x & 1 + r_A^x + r_B^y \\ 1 + r_A^y + r_B^z & 1 + r_A^y - r_B^z & 1 + r_A^y + r_B^x & 1 + r_A^y + r_B^y \pm \kappa^y \end{pmatrix}$$

Classes A through F are subdivided into families that are defined by the values of the parameters that characterize each class. Unitary transformations *do not* take a member of a family outside the family. In other words, the families are unitary invariants.

The classification presented in this section is somewhat superficial. It is based on how the state looks when expressed in terms of Pauli vectors (or Pauli quaternions, or Pauli matrices – the look and classification are the same). This should be considered only as the first step in understanding biqubit states in general.

At first glance we still cannot tell whether a given biqubit state is entangled or just a fanciful mixture that looks similar to an entangled state, but isn't.

5.10   Separability

The Englert and Metwally classification of biqubit states presented in the previous section must be supplemented with two additional conditions. The first one is obvious: every term of matrix $\boldsymbol{p}_{AB}$ must be restricted to $[0 \dots 1]$, because every term of the matrix is a probability. This imposes restrictions on the κ coefficients, together with more obvious restrictions on the $\boldsymbol{r}_{A,B}$ vectors deriving from their interepretation discovered in section 5.8, "Single qubit expectation values", page 222.

The second condition is less obvious: it may happen that a state described by $\boldsymbol{p}_{AB}$ looks perfectly OK at first glance, but is, in fact, unphysical.

An example of such a state is a Werner state of class B^+ with $\kappa = 1$, i.e.,

$$\boldsymbol{p}_{AB}^+ = \frac{1}{4} \left(\varsigma_{1A} \otimes \varsigma_{1B} + \varsigma_{xA} \otimes \varsigma_{xB} + \varsigma_{yA} \otimes \varsigma_{yB} + \varsigma_{zA} \otimes \varsigma_{zB} \right)$$

its explicit probability matrix is

$$\frac{1}{4} \begin{pmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

This matrix looks perfectly acceptable: all terms are within $[0 \ldots 1]$. But its physics looks very suspicious. What we have here is a system of two 1/2-spins that align in every direction, i.e., if we measure both spins against $\boldsymbol{e}_z$, they come out aligned. If we measure them against $-\boldsymbol{e}_z$, they come out aligned too. And this is also the case for every other direction. So this is a system of spin 1 that is spherically symmetric. But a spherically symmetric system cannot have spin 1. Only a system of spin 0 may be spherically symmetric.

> A reader who took section 2.1, "The ugly quanta", page 41, to heart may object here and say that we should not be hasty in declaring what is and what is not physical ex cathedra. That this should be decided by an experiment rather than aesthetic or even mathematical considerations. Even the most beautiful and convincing mathematics is useless if it is derived from incorrect assumptions. Well, we can say with certainty that such a state has never been observed.

If the argument about the symmetry of this state not being physical is not convincing enough, here we have another argument. The probability of a transition between this state and Bell state $\mid \Psi^- \rangle_{AB}$ is ... negative.

The Bell state $\mid \Psi^- \rangle_{AB}$ is a perfectly legitimate, pure and experimentally observed state that defines a certain direction in the biqubit Hilbert space. Its corresponding projection operator is $\mid \Psi^- \rangle_{AB} \otimes_{AB} \langle \Psi^- \mid$ and it corresponds physically to a state with spin 0. It would be perfectly OK for state $\boldsymbol{p}_{AB}^+$ to have zero probability of transition to the Bell state, but a negative probability is clearly unphysical.

Transition probability to the Bell state can be evaluated similarly to the way we did it for single-qubit states in section 4.4, "Probability amplitudes", page 126. We simply need to evaluate the following bracket, $\langle \tilde{\boldsymbol{p}}_{AB}^-, \boldsymbol{p}_{AB}^+ \rangle$:

$$\begin{aligned} \langle \tilde{\boldsymbol{p}}_{AB}^-, & \boldsymbol{p}_{AB}^+ \rangle \\ &= \Big\langle \frac{1}{4} \left(\boldsymbol{\varsigma}_A^1 \otimes \boldsymbol{\varsigma}_B^1 - \boldsymbol{\varsigma}_A^x \otimes \boldsymbol{\varsigma}_B^x - \boldsymbol{\varsigma}_A^y \otimes \boldsymbol{\varsigma}_B^y - \boldsymbol{\varsigma}_A^z \otimes \boldsymbol{\varsigma}_B^z \right), \\ & \quad \frac{1}{4} \left(\boldsymbol{\varsigma}_{1A} \otimes \boldsymbol{\varsigma}_{1B} + \boldsymbol{\varsigma}_{xA} \otimes \boldsymbol{\varsigma}_{xB} + \boldsymbol{\varsigma}_{yA} \otimes \boldsymbol{\varsigma}_{yB} + \boldsymbol{\varsigma}_{zA} \otimes \boldsymbol{\varsigma}_{zB} \right) \Big\rangle \end{aligned}$$

$$= \frac{1}{16}\left(\langle \boldsymbol{\varsigma}_A^1, \boldsymbol{\varsigma}_{1A} \rangle \langle \boldsymbol{\varsigma}_B^1, \boldsymbol{\varsigma}_{1B} \rangle - \langle \boldsymbol{\varsigma}_A^x, \boldsymbol{\varsigma}_{xA} \rangle \langle \boldsymbol{\varsigma}_B^x, \boldsymbol{\varsigma}_{xB} \rangle \right.$$

$$\left. - \langle \boldsymbol{\varsigma}_A^y, \boldsymbol{\varsigma}_{yA} \rangle \langle \boldsymbol{\varsigma}_B^y, \boldsymbol{\varsigma}_{yB} \rangle - \langle \boldsymbol{\varsigma}_A^z, \boldsymbol{\varsigma}_{zA} \rangle \langle \boldsymbol{\varsigma}_B^z, \boldsymbol{\varsigma}_{zB} \rangle \right)$$

$$= \frac{1}{16}\left(2 \times 2 - 2 \times 2 - 2 \times 2 - 2 \times 2 \right)$$

$$= -\frac{1}{2}$$

Observe that $\boldsymbol{p}_{AB}^+$ can be obtained from $\boldsymbol{p}_{AB}^-$ by replacing, for example,

$$\boldsymbol{\varsigma}_{xA} \quad \rightarrow \quad -\boldsymbol{\varsigma}_{xA}$$
$$\boldsymbol{\varsigma}_{yA} \quad \rightarrow \quad -\boldsymbol{\varsigma}_{yA}$$
$$\boldsymbol{\varsigma}_{zA} \quad \rightarrow \quad -\boldsymbol{\varsigma}_{zA} \tag{5.18}$$

while leaving all $\boldsymbol{\varsigma}_{iB}, i = x, y, z$ intact, or the other way round, but not both.

Such an operation would *not* produce a weird unphysical state if performed on a simple separable biqubit state

$$\frac{1}{2}\left(\boldsymbol{\varsigma}_{1A} + \sum_{i=x,y,z} r_A^i \boldsymbol{\varsigma}_{iA} \right) \otimes \frac{1}{2}\left(\boldsymbol{\varsigma}_{1B} + \sum_{i=x,y,z} r_B^i \boldsymbol{\varsigma}_{iB} \right)$$

because it would be equivalent to replacing $\boldsymbol{r}_A$ with $-\boldsymbol{r}_A$, and this would still be a perfectly normal, physical state made of qubit B in the same state as before and qubit A pointing in the opposite direction to the one that qubit A pointed to originally.

It would not produce a weird unphysical state if performed on a general separable biqubit state, i.e. a state that is a finite mixture of simple separable states for the same reason. It would be equivalent to replacing $\boldsymbol{r}_{Ai}$ with $-\boldsymbol{r}_{Ai}$ for every mixture component labeled by i.

Yet, when applied to an entangled state it produces an unphysical state.

It turns out [46] [29] [17] that this is a common feature of all biqubit entangled states. This amazing property was discovered by the family of Horodeckis from the University of Gdańsk in Poland and by Asher Peres from Technion in Haifa.

Why is it so? The reason for this is that an entangled biqubit can be thought of as a "new compound particle in the making", or an "old compound particle in the breaking". The latter is a more common experimental situation. Note that the making of a compound particle is the same as the breaking of a compound particle viewed backwards in time.

When a compound particle breaks the spins of the constituents must be aligned just so, in order to conserve various quantum numbers, of which angular momentum is one, and to which spin contributes. If one of the constituents gets switched the other way, artificially, the rules break *real bad* – producing, e.g., a spin-1 system that is spherically symmetric – and we end up with an unphysical configuration.

But for a biqubit that is separable, the same does not hold. Here there is no need to adhere to various conservation principles and such. The two constituents of the biqubit are fully independent and may point whichever way they wish.

Whereas the Peres-Horodeckis criterion of separability is simple and elegant, it is not all this easy to check whether a given state produced by operation (5.18) is physical or not. But we can use the bridges that lead from fiducial vectors to Pauli matrices to help ourselves in this task somewhat.

The bridge works as follows.

Let us go back to a simple separable biqubit described by

$$\boldsymbol{p}_{AB} = \boldsymbol{p}_A \otimes \boldsymbol{p}_B$$

Consider a simple energy form

$$\boldsymbol{\eta}_{AB} = \boldsymbol{\eta}_A \otimes \boldsymbol{\varsigma}_B^1 + \boldsymbol{\varsigma}_A^1 \otimes \boldsymbol{\eta}_B$$

The expectation value of energy on $\boldsymbol{p}_{AB}$ is

$$\langle \boldsymbol{\eta}_{AB}, \boldsymbol{p}_{AB} \rangle$$
$$= \langle \boldsymbol{\eta}_A, \boldsymbol{p}_A \rangle \langle \boldsymbol{\varsigma}_B^1, \boldsymbol{p}_B \rangle + \langle \boldsymbol{\varsigma}_A^1, \boldsymbol{p}_A \rangle \langle \boldsymbol{\eta}_B, \boldsymbol{p}_B \rangle$$
$$= -\boldsymbol{\mu}_A \cdot \boldsymbol{B}_A - \boldsymbol{\mu}_B \cdot \boldsymbol{B}_B$$

where $\boldsymbol{B}_A$ and $\boldsymbol{B}_B$ correspond to $\boldsymbol{B}$ in locations A and B respectively. Recall that $\langle \boldsymbol{\varsigma}^1, \boldsymbol{p} \rangle$ is 1, because there is a 1/2 in front of $\boldsymbol{\varsigma}_1$ inside $\boldsymbol{p}$.

Let us now switch to the quaternion image of the same. Here we have that

$$\boldsymbol{\rho}_{AB} = \boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B$$

and

$$\boldsymbol{H}_{AB} = \boldsymbol{H}_A \otimes \mathbf{1}_B + \mathbf{1}_A \otimes \boldsymbol{H}_B$$

Multiplying $\boldsymbol{H}_{AB}$ by $\boldsymbol{\rho}_{AB}$ yields

$$\boldsymbol{H}_{AB}\boldsymbol{\rho}_{AB}$$
$$= (\boldsymbol{H}_A \otimes \mathbf{1}_B + \mathbf{1}_A \otimes \boldsymbol{H}_B) \cdot (\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B)$$
$$= (\boldsymbol{H}_A \boldsymbol{\rho}_A) \otimes (\mathbf{1}_B \boldsymbol{\rho}_B) + (\mathbf{1}_A \boldsymbol{\rho}_A) \otimes (\boldsymbol{H}_B \boldsymbol{\rho}_B)$$

Since $\mathbf{1}\boldsymbol{\rho} = \boldsymbol{\rho}$ we can simplify this to

$$\boldsymbol{H}_{AB}\boldsymbol{\rho}_{AB} = (\boldsymbol{H}_A\boldsymbol{\rho}_A) \otimes \boldsymbol{\rho}_B + \boldsymbol{\rho}_A \otimes (\boldsymbol{H}_B\boldsymbol{\rho}_B)$$

Now recall that

$$\begin{aligned}
\boldsymbol{H}\boldsymbol{\rho} \\
&= -\frac{\mu}{2}(\boldsymbol{B}\cdot\boldsymbol{r})\,\mathbf{1} + \text{terms multiplied by sigmas} \\
&= \frac{1}{2}\langle E\rangle\,\mathbf{1} + \text{terms multiplied by sigmas}
\end{aligned}$$

so

$$\begin{aligned}
\boldsymbol{H}_{AB}\boldsymbol{\rho}_{AB} \\
&= \left(\frac{1}{2}\langle E_A\rangle\,\mathbf{1}_A + \dots\right) \otimes \boldsymbol{\rho}_B + \boldsymbol{\rho}_A \otimes \left(\frac{1}{2}\langle E_B\rangle\,\mathbf{1}_B + \dots\right) \\
&= \left(\frac{1}{2}\langle E_A\rangle\,\mathbf{1}_A + \dots\right) \otimes \left(\frac{1}{2}\mathbf{1}_B + \dots\right) + \left(\frac{1}{2}\mathbf{1}_A + \dots\right) \otimes \left(\frac{1}{2}\langle E_B\rangle\,\mathbf{1}_B + \dots\right) \\
&= \frac{1}{4}\left(\langle E_A\rangle + \langle E_B\rangle\right)\mathbf{1}_A \otimes \mathbf{1}_B + \dots
\end{aligned}$$

where "..." is a shortcut for terms that have some sigmas in them, be it one or two. So clearly, extracting what stands in front of $\mathbf{1}_A \otimes \mathbf{1}_B$ and multiplying it by 4 this time yields the right answer.

This is therefore the operation that we need here:

$$\langle E_{AB}\rangle = 4\Re_A\Re_B\left(\boldsymbol{H}_{AB}\boldsymbol{\rho}_{AB}\right) \tag{5.19}$$

Going now one step further and thinking of sigmas as matrices rather than quaternion units, we replace each $2\Re$ with its own trace operation, which yields:

$$\langle E_{AB}\rangle = \text{Tr}_A\text{Tr}_B\left(\boldsymbol{H}_{AB}\boldsymbol{\rho}_{AB}\right) \tag{5.20}$$

Various operations on tensor products of two qubits can be rewritten in the form of simple matrix and vector calculus, for example a tensor product of two 2×2 matrices can be represented by a single 4×4 matrix. This is how it works.

Consider a tensor product of two unitary vectors in a 2 dimensional Hilbert space, e.g., $|u\rangle\otimes|v\rangle$. Let

$$|u\rangle = u_0\,|0\rangle + u_1\,|1\rangle$$

and

$$|v\rangle = v_0\,|0\rangle + v_1\,|1\rangle$$

then

$$
\begin{aligned}
| \, u \rangle \otimes | \, v \rangle \; &= \; (u_0 \, | \, 0 \rangle + u_1 \, | \, 1 \rangle) \otimes (v_0 \, | \, 0 \rangle + v_1 \, | \, 1 \rangle) \\
&= \; u_0 v_0 \, | \, 0 \rangle \otimes | \, 0 \rangle + u_0 v_1 \, | \, 0 \rangle \otimes | \, 1 \rangle + u_1 v_0 \, | \, 1 \rangle \otimes | \, 0 \rangle + u_1 v_1 \, | \, 1 \rangle \otimes | \, 1 \rangle \\
&= \; u_0 v_0 \, | \, 00 \rangle + u_0 v_1 \, | \, 01 \rangle + u_1 v_0 \, | \, 10 \rangle + u_1 v_1 \, | \, 11 \rangle
\end{aligned}
$$

where we have used a shorthand $| \, 00 \rangle$ for $| \, 0 \rangle \otimes | \, 0 \rangle$ and for other combinations of the basis vectors.

But we can always re-interpret binary sequences as decimal numbers, namely:

$$
\begin{aligned}
00 \; &\equiv \; 0 \\
01 \; &\equiv \; 1 \\
10 \; &\equiv \; 2 \\
11 \; &\equiv \; 3
\end{aligned}
$$

So we can rewrite our tensor product of $| \, u \rangle \otimes | \, v \rangle$ as

$$
u_0 v_0 \, | \, \mathbf{0} \rangle + u_0 v_1 \, | \, \mathbf{1} \rangle + u_1 v_0 \, | \, \mathbf{2} \rangle + u_1 v_1 \, | \, \mathbf{3} \rangle
$$

where

$$
\begin{aligned}
| \, \mathbf{0} \rangle \; &\equiv \; | \, 0 \rangle \otimes | \, 0 \rangle \\
| \, \mathbf{1} \rangle \; &\equiv \; | \, 0 \rangle \otimes | \, 1 \rangle \\
| \, \mathbf{2} \rangle \; &\equiv \; | \, 1 \rangle \otimes | \, 0 \rangle \\
| \, \mathbf{3} \rangle \; &\equiv \; | \, 1 \rangle \otimes | \, 1 \rangle
\end{aligned}
$$

or we can also represent it in terms of a "column vector":

$$
| \, u \rangle \otimes | \, v \rangle \equiv \begin{pmatrix} x_{\mathbf{0}} \\ x_{\mathbf{1}} \\ x_{\mathbf{2}} \\ x_{\mathbf{3}} \end{pmatrix} = \begin{pmatrix} u_0 v_0 \\ u_0 v_1 \\ u_1 v_0 \\ u_1 v_1 \end{pmatrix}
$$

This can be also obtained by the following operation

$$
\begin{pmatrix} u_0 \begin{pmatrix} v_0 \\ v_1 \end{pmatrix} \\ u_1 \begin{pmatrix} v_0 \\ v_1 \end{pmatrix} \end{pmatrix}
$$

Observe the difference between the tensor product and the so called direct sum of two vectors, which would be:

$$| u \rangle \oplus | v \rangle \equiv \begin{pmatrix} u_0 \\ u_1 \\ v_0 \\ v_1 \end{pmatrix}$$

Occasionally people ask the question, "why a quantum mechanical system of two qubits is described by the tensor (direct) product and not by the direct sum of two qubits?" In classical mechanics, for example, a system of two material points is described by the direct sum of the particles' guiding vectors. The reason for this is that quantum mechanics is a probability theory and a probability of some combined outcome in a multicomponent system is a *product* of probabilities of outcomes pertaining to each component.

Now let us consider a tensor product of two 2×2 matrices. These are representations of operators acting on unitary vectors in the 2-dimensional Hilbert space, namely:

$$\begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \equiv | 0 \rangle a_{00} \langle 0 | + | 0 \rangle a_{01} \langle 1 | + | 1 \rangle a_{10} \langle 0 | + | 1 \rangle a_{11} \langle 1 | = \boldsymbol{a}$$

Let $\boldsymbol{b}$ be a similar matrix/operator. Then

$$
\begin{aligned}
\boldsymbol{a} \otimes \boldsymbol{b} &= (| 0 \rangle a_{00} \langle 0 | + | 0 \rangle a_{01} \langle 1 | + | 1 \rangle a_{10} \langle 0 | + | 1 \rangle a_{11} \langle 1 |) \\
&\quad \otimes (| 0 \rangle b_{00} \langle 0 | + | 0 \rangle b_{01} \langle 1 | + | 1 \rangle b_{10} \langle 0 | + | 1 \rangle b_{11} \langle 1 |) \\
&= | 00 \rangle a_{00} b_{00} \langle 00 | + | 00 \rangle a_{00} b_{01} \langle 01 | + | 01 \rangle a_{00} b_{10} \langle 00 | + | 01 \rangle a_{00} b_{11} \langle 01 | \\
&\quad + | 00 \rangle a_{01} b_{00} \langle 10 | + | 00 \rangle a_{01} b_{01} \langle 11 | + | 01 \rangle a_{01} b_{10} \langle 10 | + | 01 \rangle a_{01} b_{11} \langle 11 | \\
&\quad + | 10 \rangle a_{10} b_{00} \langle 00 | + | 10 \rangle a_{10} b_{01} \langle 01 | + | 11 \rangle a_{10} b_{10} \langle 00 | + | 11 \rangle a_{10} b_{11} \langle 01 | \\
&\quad + | 10 \rangle a_{11} b_{00} \langle 10 | + | 10 \rangle a_{11} b_{01} \langle 11 | + | 11 \rangle a_{11} b_{10} \langle 10 | + | 11 \rangle a_{11} b_{11} \langle 11 | \\
&= | \boldsymbol{0} \rangle a_{00} b_{00} \langle \boldsymbol{0} | + | \boldsymbol{0} \rangle a_{00} b_{01} \langle \boldsymbol{1} | + | \boldsymbol{1} \rangle a_{00} b_{10} \langle \boldsymbol{0} | + | \boldsymbol{1} \rangle a_{00} b_{11} \langle \boldsymbol{1} | \\
&\quad + | \boldsymbol{0} \rangle a_{01} b_{00} \langle \boldsymbol{2} | + | \boldsymbol{0} \rangle a_{01} b_{01} \langle \boldsymbol{3} | + | \boldsymbol{1} \rangle a_{01} b_{10} \langle \boldsymbol{2} | + | \boldsymbol{1} \rangle a_{01} b_{11} \langle \boldsymbol{3} | \\
&\quad + | \boldsymbol{2} \rangle a_{10} b_{00} \langle \boldsymbol{0} | + | \boldsymbol{2} \rangle a_{10} b_{01} \langle \boldsymbol{1} | + | \boldsymbol{3} \rangle a_{10} b_{10} \langle \boldsymbol{0} | + | \boldsymbol{3} \rangle a_{10} b_{11} \langle \boldsymbol{1} | \\
&\quad + | \boldsymbol{2} \rangle a_{11} b_{00} \langle \boldsymbol{2} | + | \boldsymbol{2} \rangle a_{11} b_{01} \langle \boldsymbol{3} | + | \boldsymbol{3} \rangle a_{11} b_{10} \langle \boldsymbol{2} | + | \boldsymbol{3} \rangle a_{11} b_{11} \langle \boldsymbol{3} | \\
&\equiv \begin{pmatrix} a_{00} b_{00} & a_{00} b_{01} & a_{01} b_{00} & a_{01} b_{01} \\ a_{00} b_{10} & a_{00} b_{11} & a_{01} b_{10} & a_{01} b_{11} \\ a_{10} b_{00} & a_{10} b_{01} & a_{11} b_{00} & a_{11} b_{01} \\ a_{10} b_{10} & a_{10} b_{11} & a_{11} b_{10} & a_{11} b_{11} \end{pmatrix}
\end{aligned}
$$

This matrix can be also obtained by the following operation:

$$
\begin{pmatrix}
a_{00} \begin{pmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{pmatrix} & a_{01} \begin{pmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{pmatrix} \\
a_{10} \begin{pmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{pmatrix} & a_{11} \begin{pmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{pmatrix}
\end{pmatrix}
$$

Now, observe that

$$
\begin{aligned}
\mathrm{Tr}_a \mathrm{Tr}_b \, \boldsymbol{a} \otimes \boldsymbol{b} \\
= \mathrm{Tr}_a \boldsymbol{a} \cdot \mathrm{Tr}_b \boldsymbol{b} = (a_{00} + a_{11}) \cdot (b_{00} + b_{11}) \\
= a_{00} b_{00} + a_{00} b_{11} + a_{11} b_{00} + a_{11} b_{11}
\end{aligned}
$$

is the same as a single trace of the large 4×4 matrix.

Matrix multiplication of the 4×4 matrix by the 4-slot vector yields another 4-slot vector, which, when contracted with a 4-slot form produces a number that is the same as would be obtained from, say, operation such as $\langle \Psi_{AB} \mid \boldsymbol{\rho}_{AB} \mid \Psi_{AB} \rangle$. So instead of working with explicit tensor products of 2×2 Pauli matrices we can switch to these 4×4 matrices instead and instead of calculating double trace operations such as $\mathrm{Tr}_A \mathrm{Tr}_B$ we can calculate single traces of the corresponding 4×4 matrices.

If the 4×4 matrix in question is made of the density quaternion $\boldsymbol{\rho}_{AB}$, then it must satisfy the same requirements that we discovered for single qubit density operators, namely:

1. It must be Hermitian.

2. Its trace must be 1.

3. It must be positive.

4. It must be idempotent if it describes a pure state.

Returning to the Peres-Horodeckis criterion for biqubit separability, we can fairly easily check that a given 4×4 density matrix of some biqubit state is *not* positive by calculating its determinant. If the determinant is negative it means that either one or three eigenvalues of the matrix are negative, which implies that on the corresponding eigenvectors the expectation values of $\boldsymbol{\rho}_{AB}$ are negative too, which, in turn, implies that the state described by $\boldsymbol{\rho}_{AB}$ is unphysical, because these expectation values are supposed to be transition probabilities.

It may happen that the 4×4 density matrix has either 2 or 4 negative eigenvalues, in which case its determinant is still positive. But the case with 4 negative

eigenvalues being negative can be easily eliminated, because in this case the trace would be negative too.

The case with 2 negative eigenvalues is harder. Here we may have to look at other matrix invariants or simply find all eigenvalues explicitly.

Let us see how this works in practice.

Appendix C lists 4×4 matrices that represent some tensor products of Pauli matrices. Let us rewrite the density operator of Bell state $\mid \Psi^- \rangle$ in the 4×4 form.

$$\rho^- = \frac{1}{4} \left(\mathbf{1} \otimes \mathbf{1} - \boldsymbol{\sigma}_x \otimes \boldsymbol{\sigma}_x - \boldsymbol{\sigma}_y \otimes \boldsymbol{\sigma}_y - \boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_z \right)$$

Making use of equations (C.1), (C.2), (C.3) and (C.4) on pages 287 and 288 results in

$$
\begin{aligned}
\rho^- \;=\; & \frac{1}{4}\left(
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
-
\begin{pmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}
\right. \\
& \left.
-
\begin{pmatrix}
0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 0
\end{pmatrix}
-
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\right) \\
\;=\; & \frac{1}{4}
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 2 & -2 & 0 \\
0 & -2 & 2 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\end{aligned}
$$

The determinant of this matrix is zero, but it is easy to see that three eigenvalues of this matrix are zero and the remaining one is $+1$. So this matrix makes a viable representation of a density operator. On no biqubit projector state $\boldsymbol{P}$ do we find that $\mathrm{Tr}\left(\boldsymbol{P}\rho^-\right) < 0$.

But now let us have a look at the Werner "state"

$$\rho^+ = \frac{1}{4} \left(\mathbf{1} \otimes \mathbf{1} + \boldsymbol{\sigma}_x \otimes \boldsymbol{\sigma}_x + \boldsymbol{\sigma}_y \otimes \boldsymbol{\sigma}_y + \boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_z \right)$$

This time the corresponding 4×4 matrix is

$$
\rho^+ \;=\; \frac{1}{4}\left(
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
+
\begin{pmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}
\right.
$$

$$+ \left(\begin{array}{cccc} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array} \right) + \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) \right)$$

$$= \frac{1}{4} \left(\begin{array}{cccc} 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{array} \right)$$

The determinant of the

$$\frac{1}{4} \left(\begin{array}{ccc} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 2 \end{array} \right)$$

sub-matrix is $-1/8$. The algebraic complement of the $1/2$ in the upper left corner of the 4×4 matrix is therefore $(-1)^{1+1}(-1/8) = -1/8$ and so, using the Laplace expansion formula the determinant of the 4×4 matrix is

$$\frac{1}{2} \cdot \left(-\frac{1}{8} \right) = -\frac{1}{16}$$

It is *negative*, which implies that either one or three eigenvalues of the 4×4 matrix are negative. Therefore the matrix does not qualify as a plausible density operator representation, because a biqubit projector state $\boldsymbol{P}$ exists such that

$$\mathrm{Tr} \left(\boldsymbol{P} \boldsymbol{\rho}^+ \right) < 0$$

It is enlightening to review the mathematical reasoning that lead the family of Horodeckis to conclude that the operation described in this section, i.e., switching the sign in front of sigmas (or varsigmas) that refer to one of the two qubits, but not both, and then checking if the resulting new state is physical, yields a necessary and sufficient condition for separability of the original state [29].

First, let us recapitulate some basic terminology. A state $\boldsymbol{\rho}_{AB}$ is called separable if it can be represented by a finite mixture of *simple* separable states, i.e., if

$$\boldsymbol{\rho}_{AB} = \sum_{i=1}^{k} p_{AB}^{i} \boldsymbol{\rho}_{iA} \otimes \boldsymbol{\rho}_{iB}$$

A state $\boldsymbol{\rho}_{AB}$ is physical if for *any* biqubit projector $\boldsymbol{P}_{AB} = |\Psi_{AB}\rangle \otimes \langle \Psi_{AB}|$

$$\mathrm{Tr} \left(\boldsymbol{P}_{AB} \boldsymbol{\rho}_{AB} \right) \geq 0$$

The density operator $\boldsymbol{\rho}_{AB}$ that satisfies this condition is also called *positive* – in this case this is a mathematical way of saying that it is *physical*.

Linear operators that act on the qubit Hilbert space themselves form a Hilbert space of their own in which a scalar product

$$\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \mathrm{Tr}\left(\boldsymbol{B}^{\dagger} \boldsymbol{A} \right)$$

can be defined, where the † operation indicates Hermitian conjugation, i.e., a matrix transposition combined with complex conjugation of the matrix terms (it does not affect any of the Pauli matrices, which are all Hermitian).

Qubit operators may be transformed into one another by the action of linear *maps*. The maps are said to be *positive* if they convert positive operators into some other positive operators. Maps are said to be *completely positive* if their tensor product with the identity is a positive map in the larger biqubit operator space. For example, if $\boldsymbol{\Lambda}_A$ is a map, then it is said to be completely positive if

$$\boldsymbol{\Lambda}_A \otimes \mathbf{1}_B$$

is positive on the space of biqubit operators.

The reasoning now runs as follows. First we observe that $\boldsymbol{\rho}_{AB}$ must be separable if and only if

$$\mathrm{Tr}\left(\boldsymbol{H}_{AB} \boldsymbol{\rho}_{AB} \right) \geq 0 \qquad\qquad (5.21)$$

for *any* Hermitian operator $\boldsymbol{H}_{AB}$ such that

$$\mathrm{Tr}\left(\boldsymbol{H}_{AB} \boldsymbol{P}_A \otimes \boldsymbol{P}_B \right) \geq 0$$

where $\boldsymbol{P}_A$ and $\boldsymbol{P}_B$ are arbitrary projectors in the space of qubits A and B respectively.

This observation seems quite obvious, at least in one direction, but there is actually a fairly simple proof that derives from certain general properties of convex spaces equipped in a scalar product, and a space of separable states has these two properties, that demonstrates the veracity of the statement in both directions, i.e., *if* and *only if* too.

A theorem proven by Jamiołkowski in 1972 is now invoked that translates the condition $\mathrm{Tr}\left(\boldsymbol{H}_{AB} \boldsymbol{\rho}_{AB} \right) \geq 0$ into the language of positive maps with a quite specific expression in place of $\boldsymbol{H}_{AB}$, namely:

$$\mathrm{Tr}\left(\boldsymbol{P}_{AB} \left(\mathbf{1}_A \otimes \boldsymbol{\Lambda}_B \boldsymbol{\rho}_{AB} \right) \right) \geq 0$$

where $\boldsymbol{P}_{AB}$ is a one-dimensional Hermitian projector in the biqubit space and $\boldsymbol{\Lambda}_B$ is an arbitrary positive map in the space of qubit B. The projector can be dropped from this condition, because it does not affect the positivity of $\mathbf{1}_A \otimes \boldsymbol{\Lambda}_B \boldsymbol{\rho}_{AB}$ and so we end up with the following theorem: $\boldsymbol{\rho}_{AB}$ *is separable if and only if* $\mathbf{1}_A \otimes \boldsymbol{\Lambda}_B \boldsymbol{\rho}_{AB}$ *is positive for* any *positive map* $\boldsymbol{\Lambda}_B$.

At this stage a theorem by Strømer and Woronowicz is invoked that says that *any* positive map $\boldsymbol{\Lambda}_B$ in two and three dimensional Hilbert spaces is of the form

$$\boldsymbol{\Lambda}_B = \boldsymbol{X}_B + \boldsymbol{Y}_B \boldsymbol{T}_B$$

where $\boldsymbol{X}_B$ and $\boldsymbol{Y}_B$ are *completely positive* maps and $\boldsymbol{T}_B$ is a matrix transposition in the space of qubit B. So

$$\begin{aligned} \mathbf{1}_A \otimes \boldsymbol{\Lambda}_B &= \mathbf{1}_A \otimes (\boldsymbol{X}_B + \boldsymbol{Y}_B \boldsymbol{T}_B) \\ &= \mathbf{1}_A \otimes \boldsymbol{X}_B + \mathbf{1}_A \otimes \boldsymbol{Y}_B \boldsymbol{T}_B \end{aligned}$$

Because $\boldsymbol{X}_B$ and $\boldsymbol{Y}_B$ are *completely positive* their tensor products with $\mathbf{1}_A$ are positive maps that do not change positivity of $\boldsymbol{\rho}_{AB}$. But $\mathbf{1}_A \otimes \boldsymbol{T}_B$ is the only term above that may change the positivity of $\boldsymbol{\rho}_{AB}$, and so we are left with the following criterion that has been literally distilled from the original formulation with $\boldsymbol{H}_{AB}$: $\boldsymbol{\rho}_{AB}$ *is separable if and only if* $\mathbf{1}_A \otimes \boldsymbol{T}_B \boldsymbol{\rho}_{AB}$ *is positive.*

The *partial transposition* operation $\mathbf{1}_A \otimes \boldsymbol{T}_B$ (it is called "partial" because it only affects one of the two qubits) does not affect $\mathbf{1}_B$, $\boldsymbol{\sigma}_{xB}$ and $\boldsymbol{\sigma}_{zB}$, but it affects $\boldsymbol{\sigma}_{yB}$, because this Pauli matrix is antisymmetric:

$$\boldsymbol{\sigma}_y^T = \left(\begin{array}{cc} 0 & -i \\ i & 0 \end{array} \right)^T = \left(\begin{array}{cc} 0 & i \\ -i & 0 \end{array} \right) = -\boldsymbol{\sigma}_y$$

Indeed, instead of replacing all three sigmas (or varsigmas) for particle A (or B, but not both) with minus sigmas (or minus varsigmas), we could have just replaced $\boldsymbol{\sigma}_y$ with $-\boldsymbol{\sigma}_y$ and already this would have converted the Bell state $\boldsymbol{p}_{AB}^-$ into an unphysical state. But why should the y direction be the only one blessed so? After all it is up to us to define *which* direction in space happens to be y.

The physical equivalent of the partial transpose operation is to reverse the polarity of one of the qubits that make a biqubit completely (all three sigmas are reversed), or to reflect its state in a mirror that is placed in the $\boldsymbol{e}_x \times \boldsymbol{e}_z$ plane (only $\boldsymbol{\sigma}_y$ is reversed.)

The partial transpose condition works *only* for 2 and 3 dimensional Hilbert spaces. For higher dimensional spaces the condition $\mathrm{Tr}\,(\mathbf{1}_A \otimes \boldsymbol{T}_B \boldsymbol{\rho}_{AB}) \geq 0$ is a *necessary* condition for $\boldsymbol{\rho}_{AB}$ to be separable, but not a sufficient one.

In summary, the Horodeckis' proof tells us that the partial transpose criterion works because it is a thorough distillation (i.e., all that's left after various irrelevant stuff is removed) of the more obvious, but much harder to apply, criterion (5.21).

The Horodeckis theorem and its physical interpretation tell us something very interesting about entangled states. They are qualitatively different from separable states and, as the result, they are very difficult to concoct. It is not enough to just bring two qubits together. All we'll end up with will be a separable state of two qubits or a mixture of such states. To produce an entangled state we usually have to do something very special, e.g., we have to split a composite particle, or we have to make an atom emit a photon. Then, only, as we saw in section 5.4, we're going to get an entangled state.

5.11   Impure quantum mechanics

Traditional unitary quantum mechanics is like Newtonian mechanics without friction. It is a highly idealized picture that ... while capturing great many phenomena very well does not entirely correspond to the real world.

In section 2.11 (pages 83 to 87), figure 2.9, page 88, we saw a vivid demonstration of the problem. There we saw that the amplitude of the observed Rabi oscillations diminished exponentially on the time scale of about a microsecond. The quantronium qubit that started in a fully polarized state, gradually lost its polarization. Our model that derived from the Schrödinger equation (via the von Neumann equation projected so as to yield evolution of qubit probabilities) could not account for this phenomenon. The $r \times B$ term ensured that only the direction and not the length of the polarization vector r would change.

Then we revisited the issue again in section 4.8.1 where we discussed a general solution to the Schrödinger equation and discovered that the equation preserved the unitarity of states being evolved, see equation (4.47) page 150.

For great many years physicists were baffled by this conundrum. How could it be that quantum physics, which explained so many phenomena so well, could not account for depolarization, dephasing, decoherence, and... ultimately the very act of measurement, the act that is at the very foundation of the unitary formalism. Yet, the Schrödinger equation, with its many variants (Pauli, Dirac), a rich assortment of Hamiltonians, and added complexities of quantum field theories and statistical physics, was the only fundamental quantum equation known to work. It was also, as we have observed in section 4.8, the *simplest* possible equation to evolve a quantum state.

Eventually the solution derived from the observation we made in section 5.8, "Single qubit expectation values". There we noticed that a *pure* biqubit maximally entangled state might look like two qubits both in completely chaotic mixed states, when the constituent qubits were measured separately, and the purity of the biqubit state asserted itself as *correlations* between the otherwise random outcomes of measurements made on the constituent qubits.

It turns out that, mathematically, every mixed quantum state can be embedded in a larger system, such that the larger system is unitary and its internal entanglement results in the observed mixed state of the embedded component. The mathematical procedure of finding such an embedding is called *purification*. So, we can state that every *impure* quantum state can be *purified*.

> This statement can be demonstrated easily for a single qubit. Recall equation (5.16) in section 5.9 that represented *every* possible *pure* biqubit

state in the following form of the D$^-$ class:

$$
\begin{aligned}
\boldsymbol{p}_{AB} \;=\; \frac{1}{4}\Bigg(&\varsigma_{1A}\otimes\varsigma_{1B} \\
&+r\left(\varsigma_{xA}\otimes\varsigma_{1B}-\varsigma_{1A}\otimes\varsigma_{xB}\right) \\
&-\left(\varsigma_{xA}\otimes\varsigma_{xB}+\sqrt{1-r^2}\left(\varsigma_{yA}\otimes\varsigma_{yB}+\varsigma_{zA}\otimes\varsigma_{zB}\right)\right)\Bigg)
\end{aligned}
$$
(5.22)

As we discovered in section 5.8 the way to eliminate the *other* qubit from a biqubit is to contract the whole biqubit with the other qubit's ς^1.

Let us extract qubit A from $\boldsymbol{p}_{AB}$ given by (5.22):

$$
\begin{aligned}
\boldsymbol{p}_A &= \langle\varsigma_B^1,\boldsymbol{p}_{AB}\rangle \\
&= \frac{1}{4}\left(2\varsigma_{1A}+2r\varsigma_{xA}\right)=\frac{1}{2}\left(\varsigma_{1A}+r\varsigma_{xA}\right)
\end{aligned}
$$

Similarly

$$
\boldsymbol{p}_B = \langle\varsigma_A^1,\boldsymbol{p}_{AB}\rangle = \frac{1}{2}\left(\varsigma_{1B}-r\varsigma_{xB}\right)
$$

They are both mixed-state qubits one pointing in the $\boldsymbol{e}_{xA}$ direction and the other one pointing in the $-\boldsymbol{e}_{xB}$ direction. But let us recall that we have obtained equation (5.22) by rotating the varsigmas pertaining to each qubit independently so as to eliminate as many terms from $\boldsymbol{p}_{AB}$ as possible. So, $\boldsymbol{e}_{xA}$ and $-\boldsymbol{e}_{xB}$ do not have to point in the same direction. Hence, any two qubits in mixed states and of the same polarization value r, pointing in arbitrary directions can be made to look like they are constituents of a pure biqubit state, if we remember to add the third term in equation (5.22), the one that describes the correlation between the qubits.

This is an algebraic, not a physical procedure. We cannot just take any two physical qubits and produce this third term,

$$
\varsigma_{xA}\otimes\varsigma_{xB}+\sqrt{1-r^2}\left(\varsigma_{yA}\otimes\varsigma_{yB}+\varsigma_{zA}\otimes\varsigma_{zB}\right),
$$

by some fancy laboratory manipulations. All that the purification theorem says is that a qubit in a mixed state can be always *thought of* as an entangled member of some larger system that is pure.

This is not to say that entangled qubits cannot be engineered. They can, but it is very difficult to do so, and even more difficult to control their entanglement for a time sufficient to carry out some meaningful computations.

Clearly, purification is not unique. The same mixed state can be made a part of various larger systems, in each case contributing to some pure state of the whole. In case of the biqubit state discussed above, we can purify qubit A by coupling

it to any of the possible qubits B pointing in various directions, as long as their polarization value r is the same as that of A. We can also purify qubit A by coupling it to larger systems, e.g., to biqubits, or tri-qubits, or to n-qubits.

So here is an idea. Why don't we treat *any* mixed state as arising from entanglement with other quantum systems? We can then use the unitary formalism and the Schrödinger equation to describe the evolution of the larger *pure* system of which the component in a mixed state is a sub-system, then we kill all the other components of the larger pure system with ς^1s, or, in the parlance of the density matrix theory, we *trace them away*, and what's left is the quantum evolution of a quantum system in a mixed state.

This may work with one snag. The evolution produced thusly will usually differ in details depending on the type of the *bath* our mixed state subsystem has been embedded in. Needless to say, a large industry exists already busying itself with such exercises. On the other hand, many common features transpire from these exercises that... seem to match quite well the observed quantum dissipative phenomena, of which qubit depolarization is but one.

By enclosing boxes within larger boxes we are led to believe that everything can be explained by the unitary formalism, until we try to embed the universe as a whole in order to explain why it is in a mixed state. We arrive at a surprising conclusion here that the universe must be in a pure state, because there is nothing out there to entangle it with, unless we invoke a multitude of other universes with which ours may be entangled, or some other multi-dimensional realms within which our universe may be only a thin membrane. This is very entertaining stuff, and it's fine as extrapolation and expoloration of ideas, but it is not exactly natural science, unless some experimental evidence is produced in support of such concepts.

Still, for a less ambitious task of modeling a depolarizing qubit, the program discussed here is workable and we're going to have a closer look at some simple and fundamental results in the following sections.

5.11.1   Nonunitary Evolution

Consider a system of two qubits, of which one is going to stand for an "environment" and the other will be subject to some evolution and observation. This is a very simplistic unitary model of a qubit interaction with its environment but it is sufficient to demonstrate a number of important non-unitary features that a quantum system entangled with its environment may display.

Let the density operator of the "environment" qubit at $t = 0$ be $\boldsymbol{\rho}_E(0)$ and the density operator of the qubit we want to measure independently of the environment

qubit be $\boldsymbol{\rho}_A(0)$ at $t = 0$. Both $\boldsymbol{\rho}_E(0)$ and $\boldsymbol{\rho}_A(0)$ are pure states, and the initial combined state is simple and separable (cf. page 191):

$$\boldsymbol{\rho}_{AE}(0) = \boldsymbol{\rho}_A(0) \otimes \boldsymbol{\rho}_E(0)$$

Suppose the biqubit is fully isolated and that its evolution is unitary, as given by equation (4.49), page 151:

$$\boldsymbol{\rho}_{AE}(t) = \boldsymbol{U}_{AE}(t)\boldsymbol{\rho}_{AE}(0)\boldsymbol{U}^\dagger_{AE}(t) = \boldsymbol{U}_{AE}(t)\left(\boldsymbol{\rho}_A(0) \otimes \boldsymbol{\rho}_E(0)\right)\boldsymbol{U}^\dagger_{AE}(t) \qquad (5.23)$$

Both qubits may become entangled with each other in the course of the evolution, and when measured individually, may appear in mixed states, but the combined system remains pure. What will the evolution of qubit A, ignoring the environment qubit B, look like?

To answer this question we rewrite equation (5.23) in more detail. Let Latin indexes, i, j, k, and l, label basis states in the Hilbert space of qubit A and let Greek indexes, α, β, γ, and δ, label basis states in the Hilber space of qubit E. Then, since $\boldsymbol{U}_{AE}$ is ultimately a linear operation, we can represent it as follows

$$\boldsymbol{U}_{AE}(t) = \sum_{i,\alpha,\beta,j} \mid i\rangle \mid \alpha\rangle U_{i\alpha\beta j}(t)\langle\beta \mid \langle j \mid,$$

where

$$U_{i\alpha\beta j} = \langle i \mid \langle\alpha \mid \boldsymbol{U}_{AE} \mid \beta\rangle \mid j\rangle \qquad (5.24)$$

Furthermore, since $\boldsymbol{\rho}_E(0)$ is pure, we can rewrite it as

$$\boldsymbol{\rho}_E(0) = \mid \Psi_E(0)\rangle\langle\Psi_E(0) \mid$$

Substituting these in equation (5.23) yields

$$\begin{aligned}\boldsymbol{\rho}_{AE}(t) \quad = \quad & \sum_{i,\alpha,\beta,j} \mid i\rangle \mid \alpha\rangle U_{i\alpha\beta j}(t)\langle\beta \mid \langle j \mid \\ & \boldsymbol{\rho}_A(0) \mid \Psi_E(0)\rangle\langle\Psi_E(0) \mid \\ & \sum_{k,\gamma,\delta,l} \mid k\rangle \mid \gamma\rangle U^\dagger_{k\gamma\delta l}(t)\langle\delta \mid \langle l \mid\end{aligned}$$

Contractions of $\mid \Psi_E(0)\rangle$ and $\langle\Psi_E(0) \mid$ with appropriate bras and kets of the Greek index type produces numbers, which multiply the $U_{i\alpha\beta j}$ and $U^\dagger_{k\gamma\delta l}$ terms resulting in

$$\begin{aligned}\boldsymbol{\rho}_{AE}(t) = \\ \sum_{i,\alpha,\beta,j,k,\gamma,\delta,l} \mid i\rangle \mid \alpha\rangle U_{i\alpha\beta j}\langle\beta \mid \Psi_E(0)\rangle\langle j \mid \boldsymbol{\rho}_A(0) \mid k\rangle\langle\Psi_E(0) \mid \gamma\rangle U^\dagger_{k\gamma\delta l}\langle\delta \mid \langle l \mid\end{aligned}$$

Now we are going to trace-out the environment qubit. This is done by taking the unsaturated dangling bras and kets of the Greek kind, here they are

$$\ldots \mid \alpha \rangle \ldots \langle \delta \mid \ldots$$

and turning them on each other:

$$\mid \alpha \rangle \langle \delta \mid \to \langle \alpha \mid \delta \rangle = \delta_{\alpha\delta},$$

where $\delta_{\alpha\delta}$ is the Kronecker delta. In expressions involving sums over all indexes, this trick produces a trace. And so we obtain

$$\boldsymbol{\rho}_A(t) =$$
$$\sum_{i,\alpha,\beta,j,k,\gamma,\delta,l} \delta_{\alpha\delta} \mid i \rangle U_{i\alpha\beta j} \langle \beta \mid \Psi_E(0) \rangle \langle j \mid \boldsymbol{\rho}_A(0) \mid k \rangle \langle \Psi_E(0) \mid \gamma \rangle U^{\dagger}_{k\gamma\delta l} \langle l \mid$$
$$= \sum_{i,\alpha,\beta,j,k,\gamma,l} \mid i \rangle U_{i\alpha\beta j} \langle \beta \mid \Psi_E(0) \rangle \langle j \mid \boldsymbol{\rho}_A(0) \mid k \rangle \langle \Psi_E(0) \mid \gamma \rangle U^{\dagger}_{k\gamma\alpha l} \langle l \mid$$
$$= \sum_{\alpha} \sum_{i,j,k,l} \mid i \rangle \left(\sum_{\beta} U_{i\alpha\beta j} \langle \beta \mid \Psi_E(0) \rangle \right) \langle j \mid \boldsymbol{\rho}_A(0) \mid k \rangle \left(\sum_{\gamma} \langle \Psi_E(0) \mid \gamma \rangle U^{\dagger}_{k\gamma\alpha l} \right) \langle l \mid$$

Expression

$$\mid i \rangle \left(\sum_{\beta} U_{i\alpha\beta j} \langle \beta \mid \Psi_E(0) \rangle \right) \langle j \mid \tag{5.25}$$

represents a matrix element of an operator, let us call it $\boldsymbol{M}_\alpha$, and what's in the brackets is $\langle i \mid \boldsymbol{M}_\alpha \mid j \rangle$, This operator acts on $\boldsymbol{\rho}_A(0)$. Similarly

$$\mid k \rangle \left(\sum_{\gamma} \langle \Psi_E(0) \mid \gamma \rangle U^{\dagger}_{k\gamma\alpha l} \right) \langle l \mid \tag{5.26}$$

represents a matrix element of $\boldsymbol{M}^{\dagger}_\alpha$. What's in the brackets is $\langle k \mid \boldsymbol{M}^{\dagger}_\alpha \mid l \rangle$. This operator acts on $\boldsymbol{\rho}_A(0)$ from the right. And so we find that

$$\boldsymbol{\rho}_A(t) = \sum_{\alpha} \boldsymbol{M}_\alpha(t) \boldsymbol{\rho}_A(0) \boldsymbol{M}^{\dagger}_\alpha(t) \tag{5.27}$$

Operations of the form given by equation (5.27) are variously called. Some people call them *quantum operations*, others call them *super operations*, and yet others just call them *maps*. They map $\boldsymbol{\rho}_A(0)$ onto $\boldsymbol{\rho}_A(t)$. They are linear, obviously, but not necessarily unitary.

We have encountered a map of a similar kind in section 5.10. The operation that reversed the polarity of one of the qubits, while keeping the other qubit intact, was a map.

Let us call a map given by (5.27) $\mathfrak{A}$ and reserve square brackets for its argument.

> The fancy looking symbol, $\mathfrak{A}$, is a Gothic "A". Mathematicians and physicists resort to Gothic letters only when they have run out of other options and are getting desperate. On the other hand, some mathematicians prefer to use normal letters and brackets for everything, because, after all, just about everything in mathematics is a mapping of some sort. But this produces formulas that can be difficult to read, because it is hard to see at first glance what's what. Our preference is for a moderately baroque notation that emphasizes geometric and transformation properties of various objects.

In our case $\mathfrak{A}$ is defined by:

$$\boldsymbol{\rho}(t) = \mathfrak{A}\left[\boldsymbol{\rho}(0)\right] = \sum_{\alpha} \boldsymbol{M}_{\alpha}(t)\boldsymbol{\rho}_A(0)\boldsymbol{M}_{\alpha}^{\dagger}(t).$$

Map $\mathfrak{A}$ must satisfy certain conditions if it is to be physical, namely, we must ensure that $\mathfrak{A}\left[\boldsymbol{\rho}(0)\right]$ produced by it is still a valid density operator.

Recall equation 5.24 above. It implies that

$$\langle i \mid \boldsymbol{M}_{\alpha} \mid j \rangle = \sum_{\beta} U_{i\alpha\beta j}\langle \beta \mid \Psi_E(0)\rangle$$

$$= \langle i \mid \langle \alpha \mid \boldsymbol{U}_{AE} \mid \Psi_E(0)\rangle \mid j \rangle,$$

which yields

$$\boldsymbol{M}_{\alpha} = \langle \alpha \mid \boldsymbol{U}_{AE} \mid \Psi_E(0)\rangle$$

This expression is a symbolic abbreviation that is often used in place of the more detailed, but less readable (5.25) and (5.26).

Now it is easy to demonstrate that

$$\sum_{\alpha} \boldsymbol{M}_{\alpha}^{\dagger}\boldsymbol{M}_{\alpha} = \mathbf{1}_A$$

This will come handy in showing that map $\mathfrak{A}$ is physical.

We begin by expanding

$$\sum_{\alpha} \boldsymbol{M}_{\alpha}^{\dagger}\boldsymbol{M}_{\alpha}$$

$$= \sum_{\alpha}\langle \Psi_E(0) \mid \boldsymbol{U}_{AE}^{\dagger} \mid \alpha\rangle\langle \alpha \mid \boldsymbol{U}_{AE} \mid \Psi_E(0)\rangle$$

$$= \langle \Psi_E(0) \mid \boldsymbol{U}_{AE}^{\dagger}\boldsymbol{U}_{AE} \mid \Psi_E(0)\rangle$$

Expression $\boldsymbol{M}_\alpha^\dagger \boldsymbol{M}_\alpha$ implies multiplication of matrices $\boldsymbol{M}$ in the (i, j) space, i.e., in the qubit A space. Additionally, the sum over $\mid \alpha \rangle \langle \alpha \mid$ produces multiplication in the (α, β) space, i.e., in the qubit E space. In effect $\boldsymbol{U}_{AE}^\dagger \boldsymbol{U}_{AE}$ is indeed full multiplication in both spaces, and therefore it must yield $\mathbf{1}_A \otimes \mathbf{1}_E$, because $\boldsymbol{U}_{AE}$ is unitary. So, we obtain

$$\sum_\alpha \boldsymbol{M}_\alpha^\dagger \boldsymbol{M}_\alpha$$
$$= \langle \Psi_E(0) \mid \mathbf{1}_A \otimes \mathbf{1}_E \mid \Psi_E(0) \rangle$$
$$= \mathbf{1}_A \langle \Psi_E(0) \mid \Psi_E(0) \rangle = \mathbf{1}_A$$

Armed with this fact we can immediately show that map $\mathfrak{A}$ preserves trace of $\boldsymbol{\rho}_A(0)$. Recall that $\mathrm{Tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{Tr}(\boldsymbol{B}\boldsymbol{A})$ and that $\mathrm{Tr}(\boldsymbol{A} + \boldsymbol{B}) = \mathrm{Tr}(\boldsymbol{A}) + \mathrm{Tr}(\boldsymbol{B})$. Hence

$$\mathrm{Tr}_A \mathfrak{A}\left[\boldsymbol{\rho}_A(0)\right] = \mathrm{Tr} \sum_\alpha \boldsymbol{M}_\alpha \boldsymbol{\rho}_A(0) \boldsymbol{M}_\alpha^\dagger$$
$$= \sum_\alpha \mathrm{Tr}\left(\boldsymbol{M}_\alpha \boldsymbol{\rho}_A(0) \boldsymbol{M}_\alpha^\dagger\right) = \sum_\alpha \mathrm{Tr}\left(\boldsymbol{\rho}_A(0) \boldsymbol{M}_\alpha \boldsymbol{M}_\alpha^\dagger\right)$$
$$= \sum_\alpha \mathrm{Tr}\left(\boldsymbol{\rho}_A(0) \boldsymbol{M}_\alpha^\dagger \boldsymbol{M}_\alpha\right) = \mathrm{Tr}\left(\sum_\alpha \boldsymbol{\rho}_A(0) \boldsymbol{M}_\alpha^\dagger \boldsymbol{M}_\alpha\right)$$
$$= \mathrm{Tr}\left(\boldsymbol{\rho}_A(0) \sum_\alpha \boldsymbol{M}_\alpha^\dagger \boldsymbol{M}_\alpha\right) = \mathrm{Tr}\left(\boldsymbol{\rho}_A(0) \mathbf{1}_A\right)$$
$$= \mathrm{Tr}\left(\boldsymbol{\rho}_A(0)\right)$$

It is easy to see that $\mathfrak{A}$ preserves positivity of $\boldsymbol{\rho}_A(0)$. Consider

$$\langle \Psi_A \mid \boldsymbol{\rho}_A(t) \mid \Psi_A \rangle$$

for an arbitrary vector $\mid \Psi_A \rangle$ in the A space. On substituting $\boldsymbol{M}_\alpha$s we get

$$\sum_\alpha \langle \Psi_A \mid \boldsymbol{M}_\alpha \boldsymbol{\rho}_A(0) \boldsymbol{M}_\alpha^\dagger \mid \Psi_A \rangle$$

But $\boldsymbol{M}_\alpha^\dagger \mid \Psi_A \rangle$ is some other vector $\mid \Phi_\alpha \rangle$ in the A space, and, since $\boldsymbol{\rho}_A(0)$ *is* positive, we find that each $\langle \Phi_\alpha \mid \boldsymbol{\rho}_A(0) \mid \Phi_\alpha \rangle$ term is positive and so their sum is positive too, which implies that $\langle \Psi_A \mid \boldsymbol{\rho}_A(t) \mid \Psi_A \rangle$ is positive.

Finally we can demonstrate that $\mathfrak{A}$ preserves hermicity of $\boldsymbol{\rho}$, i.e., that if $\boldsymbol{\rho}(0) = \boldsymbol{\rho}^\dagger(0)$ then $\boldsymbol{\rho}(t) = \boldsymbol{\rho}^\dagger(t)$:

$$\boldsymbol{\rho}^\dagger(t) \quad = \quad \left(\sum_\alpha \boldsymbol{M}_\alpha \boldsymbol{\rho}(0) \boldsymbol{M}_\alpha^\dagger\right)^\dagger$$

$$= \sum_\alpha \left(\boldsymbol{M}_\alpha \boldsymbol{\rho}(0) \boldsymbol{M}_\alpha^\dagger \right)^\dagger$$

$$= \sum_\alpha \left(\boldsymbol{M}_\alpha^\dagger \right)^\dagger \boldsymbol{\rho}^\dagger(0) \boldsymbol{M}_\alpha^\dagger$$

$$= \sum_\alpha \boldsymbol{M}_\alpha \boldsymbol{\rho}(0) \boldsymbol{M}_\alpha^\dagger$$

$$= \boldsymbol{\rho}(t),$$

where we have used $(\boldsymbol{AB})^\dagger = \boldsymbol{B}^\dagger \boldsymbol{A}^\dagger$ and $\boldsymbol{\rho}^\dagger(0) = \boldsymbol{\rho}(0)$ in the third line.

In summary, whatever $\mathfrak{A}$ does to $\boldsymbol{\rho}_A(0)$ the resulting new operator can be still interpreted as a density operator.

The *operator sum* representation of $\mathfrak{A}$ in terms of operators $\boldsymbol{M}_\alpha$, as given by equation (5.27), is not unique, because the same non-unitary evolution may result from various forms of entanglement between qubits.

5.11.2 Depolarization

Let us consider the phenomenon of depolarization that we saw in section 2.11, page 83. The unitary description of a single qubit, alone, cannot describe the gradual depolarization we observed. Schrödinger equation predicts undamped Rabi oscillations and undamped Larmor precession for such a system.

But when the single qubit is a part of a larger system, then its behavior may change dramatically.

The unitary formalism describes polarization changes in terms of rotations and flips. A typical example of a flip operation is the $\boldsymbol{\sigma}_x$ Pauli matrix, the NOT gate:

$$\boldsymbol{\sigma}_x \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a \end{pmatrix}$$

What does this operation do to vector $\boldsymbol{r}$? It swaps a and b, and therefore it changes the sign of r^z and r^y:

$$r^z = aa^* - bb^* \quad \rightarrow \quad bb^* - aa^* = -r^z$$
$$r^x = ab^* + ba^* \quad \rightarrow \quad ba^* + ab^* = r^x$$
$$r^y = \mathrm{i}(ab^* - ba^*) \quad \rightarrow \quad \mathrm{i}(ba^* - ab^*) = -r^y$$

But other flips are possible too. Let us have a look at what $\boldsymbol{\sigma}_z$ does to a qubit.

$$\boldsymbol{\sigma}_z \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ -b \end{pmatrix}$$

What is the effect of this operation on $\boldsymbol{r}$?

$$
\begin{aligned}
r^z &= aa^* - bb^* &\to& \quad aa^* - bb^* = r^z \\
r^x &= ab^* + ba^* &\to& \quad -ab^* - ba^* = -r^x \\
r^y &= \mathrm{i}(ab^* - ba^*) &\to& \quad \mathrm{i}(-ab^* + ba^*) = -r^y
\end{aligned}
$$

This operation changes the sign of r^x and r^y while leaving r^z intact.

Finally, $\boldsymbol{\sigma}_y$ does the following

$$
\boldsymbol{\sigma}_y \left(\begin{array}{c} a \\ b \end{array} \right) = \left(\begin{array}{cc} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{array} \right) \left(\begin{array}{c} a \\ b \end{array} \right) = \left(\begin{array}{c} -\mathrm{i}b \\ \mathrm{i}a \end{array} \right)
$$

And this translates into

$$
\begin{aligned}
r^z &= aa^* - bb^* &\to& \quad bb^* - aa^* = -r^z \\
r^x &= ab^* + ba^* &\to& \quad -ba^* - ab^* = -r^x \\
r^y &= \mathrm{i}(ab^* - ba^*) &\to& \quad -\mathrm{i}(ba^* - ab^*) = r^y
\end{aligned}
$$

In summary, $\boldsymbol{\sigma}_x$ rotates $\boldsymbol{r}$ by 180° around the x axis, $\boldsymbol{\sigma}_y$ rotates $\boldsymbol{r}$ by 180° around the y axis, and $\boldsymbol{\sigma}_z$ rotates $\boldsymbol{r}$ by 180° around the z axis.

Everyone of these three transformations is a unitary transformation, because

1. every Pauli matrix is Hermitian, and

2. the square of every Pauli matrix is 1

therefore for every Pauli matrix

$$
\boldsymbol{\sigma}_i \boldsymbol{\sigma}_i^\dagger = \boldsymbol{\sigma}_i \boldsymbol{\sigma}_i = \mathbf{1},
$$

which is a sufficient condition for $\boldsymbol{\sigma}_i$ to be unitary.

Suppose the qubit is a part of a larger system of three qubits with the other two qubits providing a simplistic model of an "environment". The basis states of the environment are:

$$
\begin{aligned}
\ket{0_E} \otimes \ket{0_E} &\equiv \ket{\mathbf{0}_E} \\
\ket{0_E} \otimes \ket{1_E} &\equiv \ket{\mathbf{1}_E} \\
\ket{1_E} \otimes \ket{0_E} &\equiv \ket{\mathbf{2}_E} \\
\ket{1_E} \otimes \ket{1_E} &\equiv \ket{\mathbf{3}_E}
\end{aligned}
$$

Where the notation on the right hand side is a simplified way to denote the environment biqubit basis states.

Let the initial state of the whole system be

$$| \Psi_A \rangle \otimes | \mathbf{0}_E \rangle$$

We consider a transformation $\boldsymbol{U}_{AE}$ of this system into

$$\boldsymbol{U}_{AE} | \Psi_A \rangle | \mathbf{0}_E \rangle = \sqrt{1 - \kappa} | \Psi_A \rangle \otimes | \mathbf{0}_E \rangle$$
$$+ \sqrt{\frac{\kappa}{3}} \boldsymbol{\sigma}_x | \Psi_A \rangle \otimes | \mathbf{1}_E \rangle + \sqrt{\frac{\kappa}{3}} \boldsymbol{\sigma}_y | \Psi_A \rangle \otimes | \mathbf{2}_E \rangle + \sqrt{\frac{\kappa}{3}} \boldsymbol{\sigma}_z | \Psi_A \rangle \otimes | \mathbf{3}_E \rangle,$$

where κ is the probability that the qubit is going to flip, where we allow it to flip about the x, y or z axis with equal probability of $\kappa/3$. The probability that the qubit is not going to flip is $1 - \kappa$. If the qubit does not flip, the environment stays in the $| \mathbf{0}_E \rangle$ state. If the qubit flips about the x axis, the environment switches to the $| \mathbf{1}_E \rangle$ state. If the qubit flips about the y axis, the environment switches to the $| \mathbf{2}_E \rangle$ state. And if the qubit flips about the z axis, the environment switches to the $| \mathbf{3}_E \rangle$ state. Thus, the environment responds differently to every possible flip, recording, as it were, what has happened.

The initial and final state of this operation are pure tri-qubit states, therefore the operation itself is unitary in the tri-qubit space, but it is not going to be unitary in the single qubit space. The corresponding map $\mathfrak{A}$ will have the following $\boldsymbol{M}_\alpha$ operator representation:

$$\boldsymbol{M_0} = \langle \mathbf{0}_E | \boldsymbol{U}_{AE} | \mathbf{0}_E \rangle = \sqrt{1 - \kappa} \mathbf{1}_A$$

$$\boldsymbol{M_1} = \langle \mathbf{1}_E | \boldsymbol{U}_{AE} | \mathbf{0}_E \rangle = \sqrt{\frac{\kappa}{3}} \boldsymbol{\sigma}_x$$

$$\boldsymbol{M_2} = \langle \mathbf{2}_E | \boldsymbol{U}_{AE} | \mathbf{0}_E \rangle = \sqrt{\frac{\kappa}{3}} \boldsymbol{\sigma}_y$$

$$\boldsymbol{M_3} = \langle \mathbf{3}_E | \boldsymbol{U}_{AE} | \mathbf{0}_E \rangle = \sqrt{\frac{\kappa}{3}} \boldsymbol{\sigma}_z$$

The resulting transformation of $\boldsymbol{\rho}_A$ is now

$$\boldsymbol{\rho}_A(t) = \mathfrak{A} [\boldsymbol{\rho}_A(0)]$$
$$= \boldsymbol{M_0} \boldsymbol{\rho}_A(0) \boldsymbol{M_0^\dagger} + \boldsymbol{M_1} \boldsymbol{\rho}_A(0) \boldsymbol{M_1^\dagger} + \boldsymbol{M_2} \boldsymbol{\rho}_A(0) \boldsymbol{M_2^\dagger} + \boldsymbol{M_3} \boldsymbol{\rho}_A(0) \boldsymbol{M_3^\dagger}$$
$$= (1 - \kappa) \boldsymbol{\rho}_A(0) + \frac{\kappa}{3} \boldsymbol{\sigma}_x \boldsymbol{\rho}_A(0) \boldsymbol{\sigma}_x + \frac{\kappa}{3} \boldsymbol{\sigma}_y \boldsymbol{\rho}_A(0) \boldsymbol{\sigma}_y + \frac{\kappa}{3} \boldsymbol{\sigma}_z \boldsymbol{\rho}_A(0) \boldsymbol{\sigma}_z$$

Let us apply this formula first to a general case of $\boldsymbol{\rho}_A = \frac{1}{2} (\mathbf{1} + \boldsymbol{r} \cdot \boldsymbol{\sigma})$. We can always rotate our system of coordinates so that $\boldsymbol{e}_z$ is aligned with $\boldsymbol{r}$, so without a

loss of generality we can simplify it to $\boldsymbol{\rho}_A = \frac{1}{2}\left(\mathbf{1} + r\boldsymbol{\sigma}_z\right)$. Because Pauli matrices anti-commute and square to 1 we find that

$$
\begin{aligned}
\boldsymbol{\sigma}_x\boldsymbol{\sigma}_z\boldsymbol{\sigma}_x &= -\boldsymbol{\sigma}_z\boldsymbol{\sigma}_x\boldsymbol{\sigma}_x = -\boldsymbol{\sigma}_z, \\
\boldsymbol{\sigma}_y\boldsymbol{\sigma}_z\boldsymbol{\sigma}_y &= -\boldsymbol{\sigma}_z\boldsymbol{\sigma}_y\boldsymbol{\sigma}_y = -\boldsymbol{\sigma}_z, \\
\boldsymbol{\sigma}_z\boldsymbol{\sigma}_z\boldsymbol{\sigma}_z &= \boldsymbol{\sigma}_z.
\end{aligned}
$$

And so

$$
\begin{aligned}
\mathfrak{A}\left[\boldsymbol{\rho}_A\right] &= (1-\kappa)\boldsymbol{\rho}_A + \frac{\kappa}{3}\boldsymbol{\sigma}_x\boldsymbol{\rho}_A\boldsymbol{\sigma}_x + \frac{\kappa}{3}\boldsymbol{\sigma}_y\boldsymbol{\rho}_A\boldsymbol{\sigma}_y + \frac{\kappa}{3}\boldsymbol{\sigma}_z\boldsymbol{\rho}_A\boldsymbol{\sigma}_z \\
&= \frac{1}{2}\left(\mathbf{1} + r\left(1 - \frac{4\kappa}{3}\right)\boldsymbol{\sigma}_z\right)
\end{aligned}
$$

We find that although vector $\boldsymbol{r}$ does not change its direction, it shrinks.

Suppose that $\mathfrak{A}$ is a continuous process meaning that there is a certain probability Γ of $\mathfrak{A}$ happening to the qubit per unit time. Then $\kappa = \Gamma\Delta t$. As time goes by the process repeats every Δt beginning with $t = 0$ at which time $r = r(0)$. After the first Δt, the original r shrinks to $r(\Delta t) = (1 - 4\Gamma\Delta t/3)r(0)$. After the second Δt this new r shrinks to $r(2\Delta t) = (1 - 4\Gamma\Delta t/3)r(\Delta t) = (1 - 4\Gamma\Delta t/3)(1 - 4\Gamma\Delta t/3)r(0)$.

After n such applications of $\mathfrak{A}$ the length of the polarization vector, r, will have shrunk to $r(n\Delta t) = (1 - 4\Gamma\Delta t/3)^n r(0)$.

Let $\Delta t = t/n$ then

$$
r(t) = \left(1 - \frac{4}{3}\frac{\Gamma t}{n}\right)^n r(0)
$$

The expression gets more accurate with the shrinking of Δt and with $n \to \infty$. In the limit we get

$$
r(t) = \lim_{n\to\infty}\left(1 - \frac{4}{3}\frac{\Gamma t}{n}\right)^n r(0) = \mathrm{e}^{-4\Gamma t/3}r(0),
$$

where we have explored the same trick that gave us equation (4.42) on page 148. We see that the qubit depolarizes exponentially. This is indeed what we saw in section 2.11.

5.11.3   Dephasing

Our next model is quite different. This time we are going to investigate the possible effect that entanglement with a biqubit environment E has on the third qubit A, the state of which does not change nominally. Growing entanglement, that manifests

in the environment flipping its state, is the only thing that happens. There are no random spin-flips of qubit A here.

Assuming that the initial state of the qubit-biqubit system is $\mid 0_A\rangle \mid \mathbf{0}_E\rangle$ or $\mid 1_A\rangle \mid \mathbf{0}_E\rangle$, the final state is going to be:

$$
\begin{aligned}
\boldsymbol{U}_{AE} :\mid 0_A\rangle \mid \mathbf{0}_E\rangle &\rightarrow \sqrt{1-\kappa}\mid 0_A\rangle \mid \mathbf{0}_E\rangle + \sqrt{\kappa}\mid 0_A\rangle \mid \mathbf{1}_E\rangle \\
\boldsymbol{U}_{AE} :\mid 1_A\rangle \mid \mathbf{0}_E\rangle &\rightarrow \sqrt{1-\kappa}\mid 1_A\rangle \mid \mathbf{0}_E\rangle + \sqrt{\kappa}\mid 1_A\rangle \mid \mathbf{2}_E\rangle
\end{aligned}
$$

In other words, there is a probability κ that the environment is going to flip on qubit A from $\mid \mathbf{0}_E\rangle$ to $\mid \mathbf{1}_E\rangle$ if the state of qubit A is $\mid 0_A\rangle$ and there is also similar probability κ that the environment is going to flip on qubit A from $\mid \mathbf{0}_E\rangle$ to $\mid \mathbf{2}_E\rangle$ if the state of qubit A is $\mid 1_A\rangle$. But then, there is also some probability $1-\kappa$, in both cases, that the environment is going to stay as it is.

The operator-sum representation of this interaction will have three terms, with $\boldsymbol{M}_i$ given by:

$$
\begin{aligned}
\boldsymbol{M}_0 &= \langle \mathbf{0}_E \mid \boldsymbol{U}_{AE} \mid \mathbf{0}_E\rangle \\
\boldsymbol{M}_1 &= \langle \mathbf{1}_E \mid \boldsymbol{U}_{AE} \mid \mathbf{0}_E\rangle \\
\boldsymbol{M}_2 &= \langle \mathbf{2}_E \mid \boldsymbol{U}_{AE} \mid \mathbf{0}_E\rangle,
\end{aligned}
$$

where $\boldsymbol{M}_i$ are operators that act in the space of qubit A. There is no $\boldsymbol{M}_3$ term here, because we don't make any use of $\mid \mathbf{3}_E\rangle$ in our definition of $\boldsymbol{U}_{AE}$. Qubit A entangles with a 3-dimensional subspace of biqubit E.

It is easy to see that

$$
\langle \mathbf{0}_E \mid \boldsymbol{U}_{AE} \mid \mathbf{0}_E\rangle = \sqrt{1-\kappa}\,\mathbf{1}_A
$$

For $\boldsymbol{M}_1$ we find that it acts on $\mid 0_A\rangle$ as follows

$$
\begin{aligned}
&\langle \mathbf{1}_E \mid \boldsymbol{U}_{AE} \mid 0_A\rangle \mid \mathbf{0}_E\rangle \\
&= \langle \mathbf{1}_E \mid \left(\sqrt{1-\kappa}\mid 0_A\rangle \mid \mathbf{0}_E\rangle + \sqrt{\kappa}\mid 0_A\rangle \mid \mathbf{1}_E\rangle \right) \\
&= \sqrt{\kappa}\mid 0_A,
\end{aligned}
$$

but produces zero, when acting on $\mid 1_A\rangle$. The matrix representation of $\boldsymbol{M}_1$ is therefore

$$
\boldsymbol{M}_1 = \sqrt{\kappa}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}
$$

On the other hand, $\boldsymbol{M}_2$ produces zero when acting on $\mid 0_A\rangle$, but when it acts on $\mid 1_A\rangle$ it produces

$$
\langle \mathbf{2}_E \mid \boldsymbol{U}_{AE} \mid 1_A\rangle \mid \mathbf{0}_E\rangle
$$

$$= \langle \mathbf{2}_E \mid \left(\sqrt{1 - \kappa} \mid 1_A \rangle \mid \mathbf{0}_E \rangle + \sqrt{\kappa} \mid 1_A \rangle \mid \mathbf{2}_E \rangle \right)$$
$$= \sqrt{\kappa} \mid 1_A \rangle.$$

Hence, its matrix representation is

$$\boldsymbol{M}_2 = \sqrt{\kappa} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

The resulting map $\mathfrak{A}$ is

$$
\begin{aligned}
\mathfrak{A}[\boldsymbol{\rho}_A] &= \boldsymbol{M}_0 \boldsymbol{\rho}_A \boldsymbol{M}_0^\dagger + \boldsymbol{M}_1 \boldsymbol{\rho}_A \boldsymbol{M}_1^\dagger + \boldsymbol{M}_2 \boldsymbol{\rho}_A \boldsymbol{M}_2^\dagger \\
&= (1 - \kappa)\,\boldsymbol{\rho}_A + \kappa \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \boldsymbol{\rho}_A \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \kappa \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{\rho}_A \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}
$$

Observe that

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \rho_{00} & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \rho_{11} \end{pmatrix}.$$

Hence

$$\mathfrak{A}[\boldsymbol{\rho}_A] = (1 - \kappa)\,\boldsymbol{\rho}_A + \kappa \begin{pmatrix} \rho_{00} & 0 \\ 0 & \rho_{11} \end{pmatrix} = \begin{pmatrix} \rho_{00} & (1 - \kappa)\rho_{01} \\ (1 - \kappa)\rho_{10} & \rho_{11} \end{pmatrix}$$

Now, suppose that the probability of such an entanglement happening to the qubit-biqubit system per unit time is Γ, so that $\kappa = \Gamma \Delta t$. We are going to reason here the same way we reasoned in the previous section about depolarization. After a short time Δt, the density matrix of qubit A becomes

$$\boldsymbol{\rho}_A(\Delta t) = \begin{pmatrix} \rho_{00}(0) & (1 - \Gamma \Delta t)\rho_{01}(0) \\ (1 - \Gamma \Delta t)\rho_{10}(0) & \rho_{11}(0) \end{pmatrix}$$

Then after two such time intervals:

$$\boldsymbol{\rho}_A(2\Delta t) = \begin{pmatrix} \rho_{00}(0) & (1 - \Gamma \Delta t)^2 \rho_{01}(0) \\ (1 - \Gamma \Delta t)^2 \rho_{10}(0) & \rho_{11}(0) \end{pmatrix}$$

And after n intervals such that $t = n\Delta t$ we find that

$$\boldsymbol{\rho}_A(t) = \begin{pmatrix} \rho_{00}(0) & (1 - \Gamma \Delta t)^{t/\Delta t} \rho_{01}(0) \\ (1 - \Gamma \Delta t)^{t/\Delta t} \rho_{10}(0) & \rho_{11}(0) \end{pmatrix}$$

In the limit $\Delta t \to 0$ we get

$$\lim_{\Delta t \to 0} (1 - \Gamma\Delta t) = \mathrm{e}^{-\Gamma t}$$

Thus

$$\boldsymbol{\rho}_A(t) = \begin{pmatrix} \rho_{00}(0) & \mathrm{e}^{-\Gamma t}\rho_{01}(0) \\ \mathrm{e}^{-\Gamma t}\rho_{10}(0) & \rho_{11}(0) \end{pmatrix}. \tag{5.28}$$

The exponential vanishing of the off-diagonal terms implies the exponential vanishing of x and y components of spin. Here we observe not so much depolarization, but exponentially rapid projection of the spin onto the $\boldsymbol{e}_z$ direction. This happens *not* because a force or a torque has been applied to the qubit, but because the qubit has become entangled with the biqubit that here, in this simplified model represents the environment.

This phenomenon may be thought of as a very simple unitary model of the measurement process. When a qubit, that may be polarized in any direction, encounters the measuring apparatus, it entangles with it. The effect of this entanglement is an almost instantaneous projection of the qubit onto the measurement direction of the apparatus, which is here represented by $\boldsymbol{e}_z$. Any information contained in r^x and r^y becomes lost in the process. Only information contained in r^z survives. In order to recover all information that characterizes the qubit, we have to repeat the measurements for the other two directions on the statistical ensemble of identically prepared qubits.

Recall from Section 4.8.2 (page 153) that the unitary formalism encodes information about r^x and r^y in the form of a phase difference between $|\uparrow\rangle$ and $|\downarrow\rangle$ within the *superposition*. From the unitary formalism point of view, the loss of r^x and r^y means the loss of knowledge about this phase difference, hence the term *dephasing* or *phase damping* that physicists use when discussing this process.

Figure 2.9 (B) in Chapter 2, Section 2.11, page 88, illustrates the method of measuring the decoherence time $\tau = 1/\Gamma$. In the Ramsey experiment we flip the qubit from its $|\uparrow\rangle$ state to its $|\rightarrow\rangle$ state first—in the unitary formalism, this is $(|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$. Then we let it precess about $\boldsymbol{B}_\parallel$ while the transverse buzzing field $\boldsymbol{B}_\perp$ is switched off. After some time, we turn the buzzing field $\boldsymbol{B}_\perp$ back on. If the qubit's polarization rotated around the equator of the Bloch sphere by a multiple of 2π, we are back to the starting point and the polarization continues on its march towards the south pole. If the qubit's polarization rotated around the equator of the Bloch sphere by an odd multiple of π, the qubit's polarization will move back towards the north pole. If the measurements are repeated for increasing time between the two buzzing signals we end up with "Ramsey fringes", i.e., with a curve that looks like a sinusoid.

The curve in Figure 2.9 (B) looks somewhat like a sinusoid, but it is damped. The damping here is exponential and derives from the dephasing of the qubit. The superposition state $(|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$ does not last as it rotates around the equator of the Bloch sphere. Because of the qubit's entanglement with the environment it decays exponentially towards either $|\uparrow\rangle$ or $|\downarrow\rangle$, so that when the second buzzing signal gets turned on, the polarization may no longer be on the equator of the Bloch sphere. The result of this is that as time between the two Rabi signals is extended, the qubit's final state becomes increasingly chaotic. The qubit's decoherence time τ can be read from the envelope of the curve and it can be easily seen to be on the order of a μs.

5.11.4 Spontaneous Emission

Spontaneous emission occurs when a quantum system that is in a higher energy state initially all of a sudden decays, for no apparent reason, to a lower energy state, emitting some energy quanta in the process. The quanta may be photons, but they may be other particles, too.

A reverse process to spontaneous emission is spontaneous absorption. Here a quantum system absorbs energy quanta from its environment and upgrades itself to a higher energy state.

Both processes are different from unitary absorption or emission of energy as described by the Schrödinger equation. The difference is that here we don't have an obvious driver. We don't know the reasons for the decay or for absorption. There may not be any reasons, or, as we shall see shortly, the reason may be the entanglement with the environment. In both cases, the resulting process is non-unitary, and described in terms of map $\mathfrak{A}$.

The simplest possible model of spontaneous emission is given by the following two equations that define $\boldsymbol{U}_{AE}$.

$$\boldsymbol{U}_{AE} \,|\, 0_A \rangle \,|\, \boldsymbol{0}_E \rangle \;\; = \;\; |\, 0_A \rangle \,|\, \boldsymbol{0}_E \rangle \tag{5.29}$$

$$\boldsymbol{U}_{AE} \,|\, 1_A \rangle \,|\, \boldsymbol{0}_E \rangle \;\; = \;\; \sqrt{1-\kappa} \,|\, 1_A \rangle \,|\, \boldsymbol{0}_E \rangle + \sqrt{\kappa} \,|\, 0_A \rangle \,|\, \boldsymbol{1}_E \rangle \tag{5.30}$$

In plain English, if the observed qubit A is in the ground state, it stays in the ground state. Nothing changes. But if it is in the higher energy state, $|\, 1_A \rangle$, then there is a probability of κ that it is going to decay, transferring the energy to the environment that now flips from $|\, \boldsymbol{0}_E \rangle$ to $|\, \boldsymbol{1}_E \rangle$. And then, there is also a probability $1 - \kappa$ that qubit A is going to stay in state $|\, 1_A \rangle$.

This process differs from depolarization. If this was depolarization, then we would also allow state $|\, 0_A \rangle$ to flip to $|\, 1_A \rangle$. We use only two states of the environment

here: $|\,\boldsymbol{0}_E\rangle$ and $|\,\boldsymbol{1}_E\rangle$, so we are going to have only two operators $\boldsymbol{M}_i$ with $i = 0$ and 1. They are defined by

$$\boldsymbol{M}_0 \;=\; \langle\boldsymbol{0}_E\,|\,\boldsymbol{U}_{AE}\,|\,\boldsymbol{0}_E\rangle \tag{5.31}$$

$$\boldsymbol{M}_1 \;=\; \langle\boldsymbol{1}_E\,|\,\boldsymbol{U}_{AE}\,|\,\boldsymbol{0}_E\rangle. \tag{5.32}$$

In order to reconstruct the matrices of $\boldsymbol{M}_0$ and $\boldsymbol{M}_1$ in the space of qubit A we need to find out what are the effects of $\boldsymbol{M}_0$ and $\boldsymbol{M}_1$ acting on the basis vectors of qubit A. And so

$$\boldsymbol{M}_0\,|\,0_A\rangle \;=\; \langle\boldsymbol{0}_E\,|\,\boldsymbol{U}_{AE}\,|\,0_A\rangle\,|\,\boldsymbol{0}_E\rangle = \langle\boldsymbol{0}_E\,(|\,0_A\rangle\,|\,\boldsymbol{0}_E\rangle) = |\,0_A\rangle \tag{5.33}$$

$$\boldsymbol{M}_0\,|\,1_A\rangle \;=\; \langle\boldsymbol{0}_E\,|\,\boldsymbol{U}_{AE}\,|\,1_A\rangle\,|\,\boldsymbol{0}_E\rangle = \langle\boldsymbol{0}_E\,\left(\sqrt{1-\kappa}\,|\,1_A\rangle\,|\,\boldsymbol{0}_E\rangle + \sqrt{\kappa}\,|\,0_A\rangle\,|\,\boldsymbol{1}_E\rangle\right)$$

$$\;=\; \langle\boldsymbol{0}_E\sqrt{1-\kappa}\,|\,1_A\rangle\,|\,\boldsymbol{0}_E\rangle = \sqrt{1-\kappa}\,|\,1_A\rangle \tag{5.34}$$

This tells us that the matrix of $\boldsymbol{M}_0$ is diagonal and looks as follows:

$$\boldsymbol{M}_0 = \left(\begin{array}{cc} 1 & 0 \\ 0 & \sqrt{1-\kappa} \end{array}\right). \tag{5.35}$$

Similarly, for $\boldsymbol{M}_1$ we find:

$$\boldsymbol{M}_1\,|\,0_A\rangle \;=\; \langle\boldsymbol{1}_E\,|\,\boldsymbol{U}_{AE}\,|\,0_a\rangle\,|\,\boldsymbol{0}_E\rangle = \langle\boldsymbol{1}_E\,|\,(|\,0_A\rangle\,|\,\boldsymbol{0}_E\rangle) = 0 \tag{5.36}$$

$$\boldsymbol{M}_1\,|\,1_A\rangle \;=\; \langle\boldsymbol{1}_E\,|\,\boldsymbol{U}_{AE}\,|\,1_a\rangle\,|\,\boldsymbol{0}_E\rangle = \langle\boldsymbol{1}_E\,|\,\left(\sqrt{1-\kappa}\,|\,1_A\rangle\,|\,\boldsymbol{0}_E\rangle + \sqrt{\kappa}\,|\,0_A\rangle\,|\,\boldsymbol{1}_E\rangle\right)$$

$$\;=\; \langle\boldsymbol{1}_E\,|\,\sqrt{\kappa}\,|\,0_A\rangle\,|\,\boldsymbol{1}_E\rangle = \sqrt{\kappa}\,|\,0_A\rangle \tag{5.37}$$

Hence, the resulting matrix of $\boldsymbol{M}_1$ is

$$\boldsymbol{M}_1 = \left(\begin{array}{cc} 0 & \sqrt{\kappa} \\ 0 & 0 \end{array}\right) \tag{5.38}$$

And the map $\mathfrak{A}$ looks as follows:

$$\mathfrak{A}\,[\boldsymbol{\rho}_A] = \boldsymbol{M}_0\boldsymbol{\rho}_A\boldsymbol{M}_0^\dagger + \boldsymbol{M}_1\boldsymbol{\rho}_A\boldsymbol{M}_1^\dagger$$

$$= \left(\begin{array}{cc} 1 & 0 \\ 0 & \sqrt{1-\kappa} \end{array}\right)\boldsymbol{\rho}_A\left(\begin{array}{cc} 1 & 0 \\ 0 & \sqrt{1-\kappa} \end{array}\right) + \left(\begin{array}{cc} 0 & \sqrt{\kappa} \\ 0 & 0 \end{array}\right)\boldsymbol{\rho}_A\left(\begin{array}{cc} 0 & 0 \\ \sqrt{\kappa} & 0 \end{array}\right)$$

For a general $\boldsymbol{\rho}_A = \left(\begin{array}{cc} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{array}\right)$ this becomes

$$\mathfrak{A}\,[\boldsymbol{\rho}_A] \;=\; \left(\begin{array}{cc} \rho_{00} & \sqrt{1-\kappa}\rho_{01} \\ \sqrt{1-\kappa}\rho_{10} & (1-\kappa)\rho_{11} \end{array}\right) + \left(\begin{array}{cc} \kappa\rho_{11} & 0 \\ 0 & 0 \end{array}\right)$$

$$\;=\; \left(\begin{array}{cc} \rho_{00} + \kappa\rho_{11} & \sqrt{1-\kappa}\rho_{01} \\ \sqrt{1-\kappa}\rho_{10} & (1-\kappa)\rho_{11} \end{array}\right) \tag{5.39}$$

Now, suppose we apply $\mathfrak{A}$ twice. The resulting new $\boldsymbol{\rho}$ in terms of the original $\boldsymbol{\rho}$ is going to look as follows.

$$\mathfrak{A}\left[\mathfrak{A}\left[\boldsymbol{\rho}_A\right]\right] = \begin{pmatrix} \rho_{00} + \kappa\rho_{11} + \kappa(1-\kappa)\rho_{11} & \sqrt{1-\kappa}\sqrt{1-\kappa}\rho_{01} \\ \sqrt{1-\kappa}\sqrt{1-\kappa}\rho_{10} & (1-\kappa)(1-\kappa)\rho_{11} \end{pmatrix} \qquad (5.40)$$

As we apply $\mathfrak{A}$ n times it is easy to see that the off-diagonal terms simply get multiplied by $(1-\kappa)^{n/2}$ and the ρ_{11} term gets multiplied by $(1-\kappa)^n$. The ρ_{00} term behaves as follows:

$$\begin{aligned} \text{for } n = 0: \quad & \rho_{00} \\ \text{for } n = 1: \quad & \rho_{00} + \kappa\rho_{11} \\ \text{for } n = 2: \quad & \rho_{00} + \kappa\rho_{11} + \kappa(1-\kappa)\rho_{11} \\ \text{for } n = 3: \quad & \rho_{00} + \kappa\rho_{11} + \kappa(1-\kappa)\rho_{11} + \kappa(1-\kappa)^2\rho_{11} \\ & \cdots \end{aligned}$$

So for an arbitrary n this becomes

$$\rho_{00} + \kappa\rho_{11} + \kappa(1-\kappa)\rho_{11} + \ldots + \kappa(1-\kappa)^{n-1}\rho_{11}$$

$$= \rho_{00} + \kappa\rho_{11}\sum_{k=0}^{n-1}(1-\kappa)^k$$

The sum in this expression is the sum of a geometric series, which evaluates to

$$\sum_{k=0}^{n-1}(1-\kappa)^k = \frac{1-(1-\kappa)^n}{1-(1-\kappa)} = \frac{1-(1-\kappa)^n}{\kappa} \qquad (5.41)$$

So,

$$\kappa\sum_{k=0}^{n-1}(1-\kappa)^k = 1-(1-\kappa)^n \qquad (5.42)$$

And so the $(0,0)$ term of the density matrix becomes

$$\rho_{00} + \rho_{11}\left(1-(1-\kappa)^n\right) \qquad (5.43)$$

Now, let us again exploit the trick that has served us so well in the previous sections. We assume that $\kappa = \Gamma\Delta t$, where $n\Delta t = t$. Then

$$\mathfrak{A}^n\left[\boldsymbol{\rho}_A\right] = \begin{pmatrix} \rho_{00} + \rho_{11}\left(1-\left(1-\Gamma\frac{t}{n}\right)^n\right) & \left(1-\Gamma\frac{t}{n}\right)^{n/2}\rho_{01} \\ \left(1-\Gamma\frac{t}{n}\right)^{n/2}\rho_{10} & \left(1-\Gamma\frac{t}{n}\right)^n\rho_{11} \end{pmatrix} \qquad (5.44)$$

In the limit $n \to \infty$ and $\Delta t \to 0$ such that $n\Delta t = t$ we obtain

$$\boldsymbol{\rho}(t) = \begin{pmatrix} \rho_{00}(0) + \rho_{11}(0)\left(1 - \mathrm{e}^{-\Gamma t}\right) & \mathrm{e}^{-\Gamma t/2}\rho_{01}(0) \\ \mathrm{e}^{-\Gamma t/2}\rho_{10}(0) & \mathrm{e}^{-\Gamma t}\rho_{11}(0) \end{pmatrix} \tag{5.45}$$

We find that as time flows the system converges exponentially on a state that "points up", which is the lower energy state, $|\,0_A\rangle$.

If the original state $|\,\Psi(0)\rangle$ had a non-vanishing amplitude of being registered in the $|\,1_A\rangle$ state, the amplitude would decay exponentially to zero, as would amplitudes of finding the system in any of the transverse states, $|\!\rightarrow\rangle$ or $|\otimes\rangle$. For this reason physicists call this process *amplitude damping.*

Spontanous emission has one beneficial side effect. We can use it to force a quantum system, which may have been in some thermal chaotic state initially, to cool down and become pure. This is how quantum computations often begin.

5.12   Schrödinger's Cat

Here is the sad story of the Schrödinger's cat.

Schrödinger conceived of a following *gedanken* experiment.

A quantum system, for example, an atom of Rubidium is put in a superposition state of two so called *circular Rydberg* levels with principal quantum numbers of 51 (we call this the $|\,e\rangle$ state) and 50 (we call this the $|\,g\rangle$ state).

> Circular Rydberg states are states of multi-electron atoms, for example atoms such as Rubidium, which has 37 electrons normally. They are characterized by large magnetic quantum numbers that derive from the orbital motion of many electrons around the nucleus. They are also characterized by large principal quantum numbers of their valence electrons, i.e., electrons that are in the outer shells of the atom. They tend to respond strongly to electromagnetic stimulation and have long lifetimes. The circular Rydberg states of Rubidium mentioned here have a lifetime of 30 ms. For these reasons multi-electron atoms in circular Rydberg states are often used in quantum experiments, including quantum computing systems. The other reason is that the energy difference between $|\,e\rangle$ and $|\,g\rangle$ is in the microwave range, 51.099 GHz. Electromagnetic radiation in this range can be controlled with great precision.

The atom is observed by a detector, which is connected to a vial filled with poisonous gas. The detector works by smashing the vial if it detects the atom in the $|\,e\rangle$ state and not smashing it if it detects the atom in the $|\,g\rangle$ state.

Now comes the cruelty of this needless experiment—*gedanken* or not. The contraption is put in a sealed box together with a live cat. And Dr. Schrödinger says, "because the cat becomes coupled to the atom through the contraption with the vial of poison, the amount of poison being sufficient to kill the cat, the cat, like the atom, should be in a superposition of two states: $|\text{alive}\rangle$ and $|\text{dead}\rangle$. And yet, if we were to peek into the box, we would find the cat *either* alive *or* dead, and not in between."

The real purpose of this somewhat exaggerated *gedanken* experiment was to point out that we could not separate the logic of the microscopic, quantum world, from the logic of the macroscopic world—that we could always conceive of some situation in which the states of the microscopic world had a direct bearing on macroscopic systems.

As is often the case with *gedanken* experiments, the conclusions drawn depend on great many untold assumptions. *Gedanken* experiments became fashionable after Einstein's initial success with them. But Einstein's experiments dealt with macroscopic systems and with macroscopic physics, all of which we are familiar with. The situations were made strikingly simple so as to expose a particular point Einstein was after. But there is nothing so simple when it comes to coupling between an atom and a macroscopic system. The difference between the two is described by the Avogadro number, 6.02252×10^{23}/mol—there is an Avogadro number of interacting atoms in the detector versus only one atom the detector observes.

We have seen in this chapter the degree of complexity that arises when we move from contemplating one qubit to contemplating two qubits. Imagine the mind-boggling complexity if we were to switch from one qubit to Avogadro number of qubits. What would the density operator of such a system look like? What would be the atom's evolution if we were to trace out the detector? And what would be the detector's evolution if we were to trace out the atom?

As to the cat, we should really take it out of the box, because its presence only confuses the issue. The difference between $|\text{live}\rangle$ and $|\text{dead}\rangle$ is nowhere near as sharply defined as the difference between the $|e\rangle$ and $|g\rangle$ states of the atom. The cat may be sick, very sick, half dead, barely alive, or... very angry. It may also rub against the detector and trigger the mechanism regardless of the state of the atom.

Rather than focusing on the cat, we could focus on the vial itself and contemplate a superposition of $|\text{smashed}\rangle$ and $|\text{whole}\rangle$. And we don't need to fill it with poison.

But the presence of the vial confuses the issue just as much as the presence of the cat. Clearly, it is the detector itself and its interaction with the atom that is of interest to us here. How the detector manifests the detected state of the atom, be it by smashing a vial, or beeping, or moving a pointer, is of secondary impor-

tance. What is important, is that the detector entangles with the observed atom somehow—some physical interaction must be present—and eventually responds to the state of the atom by counting either "up" or "down", with the "up" response on the detector side leaving the atom in the $|\,e\rangle$ state and the "down" response on the detector side leaving the atom in the $|\,g\rangle$ state. The atom will *dephase,* and the detector will respond to the atom's dephasing by dephasing itself into a state that can be read by us.

This is how detectors work.

What greatly puzzled physicists in Schrödinger's days was that back then they believed that it was the "act of observation", however ill-defined, that was responsible for the observed quantum system flipping to one of its basis states. Some even believed that a conscious observer was needed. This resulted in one of the axioms of quantum mechanics stating that "upon observation a quantum system finds itself irreversibly in one of its basis states". It is for this reason that they also believed that the cat would be fine as long as we wouldn't peek into the box. In other words that it would be our act of observing the cat that would throw it out of the superposition of $\left(|\,\text{dead}\rangle+|\,\text{alive}\rangle\right)/\sqrt{2}$.

The other issue here is what we referred to in Section 4.3, page 117, as "The Superstition of Superposition". There we argued that a state that is in a superposition of two basis states is not in both states at the same time, but, instead, it is in neither. It is in a third state that is altogether different. The physical meaning of this third state is not always easy to figure out, but it can be always identified by evaluating the full density operator and the full vector of probabilities for the state. Consequently, even if we could put the cat in the superposition of $|\,\text{dead}\rangle$ and $|\,\text{alive}\rangle$, it may not necessarily be something out of this world—it may, instead, be just $|\,\text{sick}\rangle$.

Today we have a more sophisticated view of what goes on. We don't need an axiom. We can derive the dynamics of the combined system from the basic principles of quantum mechanics using just the Schrödinger equation, in principle and in practice even, as long as the detector itself is not too large. A detector comprising the Avogadro number of atoms is out of the question here. It is not computable. But a detector comprising no more than 10 qubits can be analyzed fully and if the corresponding system can be implemented in a laboratory, the theoretical predictions can be compared against the actual measurements.

5.12.1   The Haroche–Ramsey Experiment

Just such an experiment was carried out by Haroche and his colaborators, Brune, Hagley, Dreyer, Maître, Maali, Wunderlich and Raimond, at the Kastler Brossel Laboratory of the Paris l'Ecole Normale Supérieure in 1996 [9]. The experiment is similar to Ramsey measurement, but with an important modification, hence the name.

Haroche's group created an "atom + measuring apparatus" system in which the measuring apparatus was a mescoscopic cavity holding up to 10 photons. They demonstrated that the photons in the cavity themselves were put in a superposition of states by their interaction with the atom. Then they observed their dephasing and transformation of the cavity state into a statistical mixture. The observed behaviour of the system was contrasted with theoretical predictions [12] obtained by an analysis somewhat like what we have presented in section 5.11—with the difference that instead of entangling a qubit with another qubit, here we entangle it with 10 photons—and a highly accurate match was demonstrated.
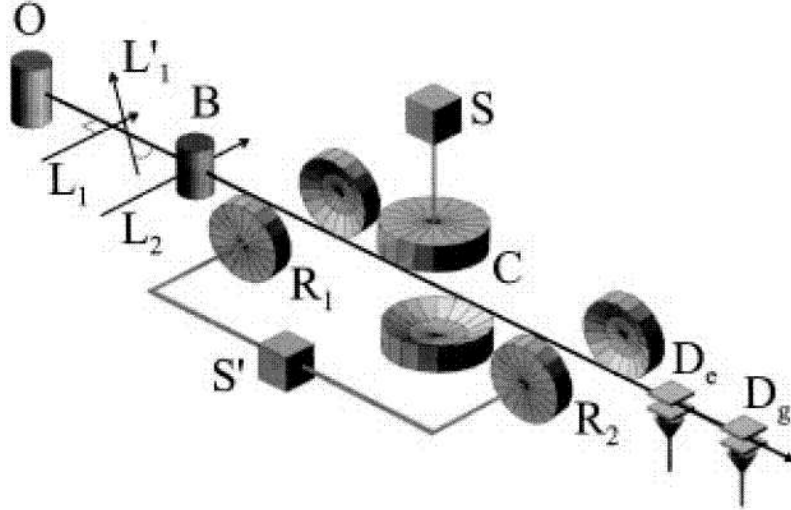
Figure 5.9: Setup for the Haroche-Ramsey experiment, from [9]

The measurement setup for the experiment is shown in Figure 5.9. The whole

apparatus is placed in vacuum and cooled to $0.6\,\mathrm{K}$ to reduce thermal radiation to a negligible level.

The cylindrical can labeled O is an oven that effuses Rubidium atoms. The atoms emerging from the oven are first conditioned by diode lasers L_1 and L'_1 so that a certain subset of them, namely the ones that move with velocity of $400\pm6\,\mathrm{m/s}$, and these atoms only, are in a state that is then pumped into a circular Rydberg state $|\,e\rangle$ in the box labeled B. This procedure prepares on average $1/2$ an atom every $1.5\,\mathrm{ms}$, and it takes about $2\,\mu\mathrm{s}$ in B to condition the atom. Other atoms will be naturally filtered away by the remaining part of the experiment, so we can forget about them.

As the selected atoms cross the cavity labeled R_1, in which they spend a precise amount of time, on account of their selected velocity, their quantum state $|\,e\rangle$ is rotated by $\pi/2$ into $(|\,e\rangle + |\,g\rangle)/\sqrt{2}$. This is the Rabi rotation we had studied in section 2.10, page 75.

Let us, for the moment, assume that there is no cavity labeled S. As the atoms fly from the microwave cavity R_1 to the microwave cavity R_2 they precess, and are then rotated again by $\pi/2$ in R_2. So, we end up with a pure Ramsey experiment, like the one discussed in section 2.10, that is here carried out on atoms of Rubidium in these two circular Rydberg states. The atoms have more states than just $|\,e\rangle$ and $|\,g\rangle$, but the dynamics of the experiment is confined to these two states only. Consequently, this is a qubit experiment, which is good, because qubits are simple and easy to understand.

Past the microwave cavity R_2 there are two field ionization detectors, D_e and D_g. The first one is tuned to detect atoms in state $|\,e\rangle$ and the second one is tuned to detect atoms in state $|\,g\rangle$. Both have detection efficiency of $40\pm15\%$.

The frequency of the Rabi oscillations field in cavities R_1 and R_2 is varied slightly—by up to $10\,\mathrm{kHz}$—around the resonance Rabi frequency for the $|\,e\rangle \rightarrow |\,g\rangle$ transition, $\nu_0 = 51.099\,\mathrm{GHz}$. The variation is only two parts per 10 million, sic!. This has a similar effect to stretching or shrinking the free precession time between the two Ramsey pulses, so that when the atoms are finally detected either at D_e or D_g we observe Ramsey fringes as shown in Figure 5.10 (a). It's just that instead of observing the fringes in function of time elapsed between the two Ramsey pulses, here we observe them in function of the pulse frequency.
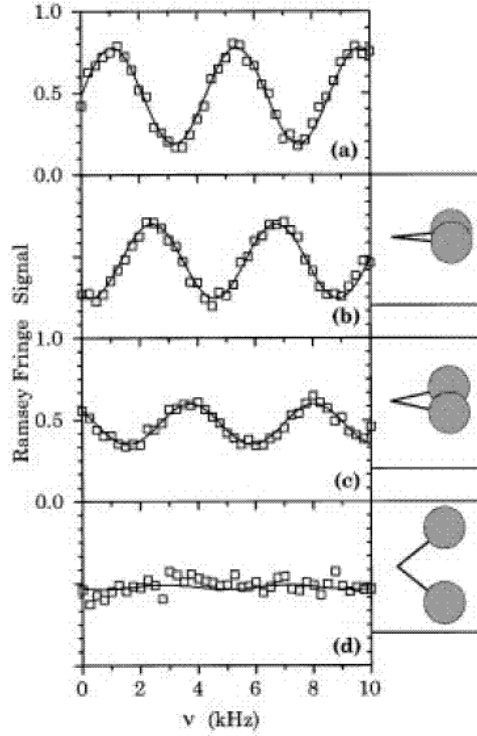
Figure 5.10: Ramsey fringes, from [9]. (a) the microwave cavity C contains no photons, (b)–(d) the microwave cavity C contains photons, frequency detuning is (b) 712 kHz, (c) 347 kHz, (d) 104 kHz.

What is plotted in the diagrams of Figure 5.10 is the measured probability of detecting an atom in the $\mid g \rangle$ state, based on the exploration of a statistical ensemble of 1000 events per each point of the graph and sampling for 50 discrete values of frequency ν [27]. To collect enough statistics for each graph takes about 10 minutes. The expected standard deviation for a count of 1000 is $\sqrt{1000} \approx 32$ (see, e.g., [19]), which yields the relative error in the estimated probability of about 3%. The smooth lines are sinusoids fitted through the experimental data.

Even though the detector efficiencies are about 40% only, here we take a ratio of N_g to $N_g + N_e$, where N_g and N_e are the actually registered counts, so the curve

in Figure 5.10 should vary between 0 and 1 in principle. Instead it is squashed to between 0.22 and 0.78, on average, or, to put it in another way, its *contrast* is reduced to 55±5%. This is, amongst other reasons, because of static and microwave field inhomogeneities over the diameter of the qubit beam, which is 0.7 mm. The finite lifetime of the $|e\rangle$ and $|g\rangle$ states contributes to this loss of contrast as well.

Now let us turn our attention to the cavity called C. C stands for "cat". This is the mesoscopic "detector" the atom interacts with. The cavity is made of two concave superconducting niobium mirrors separated by about 2.7 cm. The diameter of each mirror is 5 cm and the curvature radius of their inner surface is 4 cm. The electromagnetic fields trapped between the mirrors are focused on a small region between them that is about 6 mm across. The Rubidium beam traverses the region in about 19 μs.

The field in the cavity is quantized and coherent. This had been demonstrated prior to the Haroche–Ramsey experiment by the the Haroche group [10] and was one of the first such observations in history, even though the idea of the electromagnetic field quantization goes back to the Einstein's photoelectric effect paper of 1905 [15].[7] The cavity field quantization and coherence can be observed because it is extremely weak and the cavity is cooled to near absolute zero. The average number of photons in the cavity varies between 0 and 10, their lifetime is about 160 μs and their collective state prior to the interaction with the atom can be described symbolically by $|\alpha\rangle$.

Interaction between the qubit and this kind of a quantum field is not quite like interaction between the qubit and the *classical* Rabi field, which we discussed in Section 2.10. Here the interaction is with just a handful of photons, the number of which is quantumly uncertain, meaning that there is a certain distribution $P(n)$ that gives us the probability of there being n photons in the cavity. This problem was worked out by Jaynes and Cummings in 1963 [32] and the result is such that if the cavity is filled with photons of the $|e\rangle \rightarrow |g\rangle$ transition frequency, then the probability of the transition is

$$P_{|e\rangle\rightarrow|g\rangle}(t) = \sum_n P(n)\sin^2\left(\Omega\sqrt{n+1}t\right) \tag{5.46}$$

where t is the time of interaction and $\Omega = 2\pi \times 24$ kHz is the qubit-field coupling. This parameter plays a role similar to the Rabi frequency in this special context.

The exact distance between the mirrors can be varied, which results in tuning the cavity. It can be tuned so that the field frequency in it matches the $|e\rangle \rightarrow |g\rangle$

---

[7]In the Planck's original derivation of the black body radiation formula, it was the matter's ability to absorb and emit radiation that was quantized, not the electromagnetic field.

transition frequency, but instead it is *detuned* by between $\delta = 2\pi \times 70\,\text{kHz}$ and $\delta = 2\pi \times 800\,\text{kHz}$. In effect, when the Rubidium atoms traverse the field region, the frequency of the field is too far from the Rabi frequency and the atoms spend too little time in the field region for any energy exchange to take place between them. Instead the field in the cavity becomes coupled to the atom, which changes the phase of the field by

$$\phi = \frac{\Omega^2 t}{\delta} \quad \text{for} \quad \Omega << \delta$$

if the atom was in the $\mid e \rangle$ state prior to entering the cavity and $-\phi$ if the atom was in the $\mid g \rangle$ state.

This is how the field in the cavity C becomes the "detector" of the atom's state.

After entanglement with the atom, the combined state of the photons+atom system becomes $\mid e \rangle \otimes \mid \alpha e^{i\phi} \rangle$ or $\mid g \rangle \otimes \mid \alpha e^{-i\phi} \rangle$ and if the atom was in the superposition state $(\mid e \rangle + \mid g \rangle) / \sqrt{2}$ prior to entering the cavity C, the combined state of the photons+atom system, after $19\,\mu\text{s}$ of the interaction in the cavity, ends up in the superposition

$$\frac{1}{\sqrt{2}} \left(\mid e \rangle \otimes \mid \alpha e^{i\phi} \rangle + \mid g \rangle \otimes \mid \alpha e^{-i\phi} \rangle \right).$$

This has a profound effect on the Ramsey fringes that are detected by D_g.

Figure 5.10 (b) shows Ramsey fringes when the cavity C is filled with 9.5 photons on average and it is detuned from the Rabi frequency by $712\,\text{kHz}$. We can see that the contrast has diminished somewhat and the peaks have shifed to the right a little. The insert on the right hand side of the graph shows the field phase shift ϕ, as phasors, for both the $\mid e \rangle$ and the $\mid g \rangle$ states.

Figure 5.10 (c) shows the same, but this time the cavity is detuned by $347\,\text{kHz}$. We are getting closer to the Rabi frequency, thus increasing the entanglement between the atom and the photons in the cavity. The response registered by D_g shows even more diminished contrast and even more shift in the location of the peaks. The accompanying cavity field phase shift ϕ is somewhat larger than in the previous case.

Finally, Figure 5.10 (d) shows Ramsey fringes for the cavity detuning of $104\,\text{kHz}$. This time we are quite close to the Rabi frequency, though still not close enough for the full interaction described by equation (5.46) to take place. The entanglement between the atom and the photons is stronger still and the Ramsey fringes almost disappear. The cavity field phase shift ϕ is larger still.

The blue balls at the tips of the phasors in the inserts represent the uncertainty in the field's phase, that is due to quantum fluctuations of the field. We can see that

in Figure 5.10 (d) the phase difference between the two field states is sufficiently large to be resolved even in the presence of the uncertainties. On the other hand, the separation in Figure 5.10 (b) is too small and the two states are not resolvable.

The connection between our ability to resolve the two states of the "cat", $\mid \alpha e^{i\phi}\rangle$ and $\mid \alpha e^{-i\phi}\rangle$, and the disappearance of the fringe pattern, which in the unitary picture can be thought of as resulting from the interference of two amplitudes

$$\langle g \mid e \rangle = \langle g \mid R_1 \mid e \rangle + \langle g \mid R_2 \mid e \rangle$$

is characteristic of quantum physics and shows up in many other situations. It is often explained by hand-waving arguments about the conspiracy of nature. However, here we can carry out detailed calculations, because the problem of coupling a qubit to 10 photons is still computable, and compare theoretical predictions to observed, experimental data.
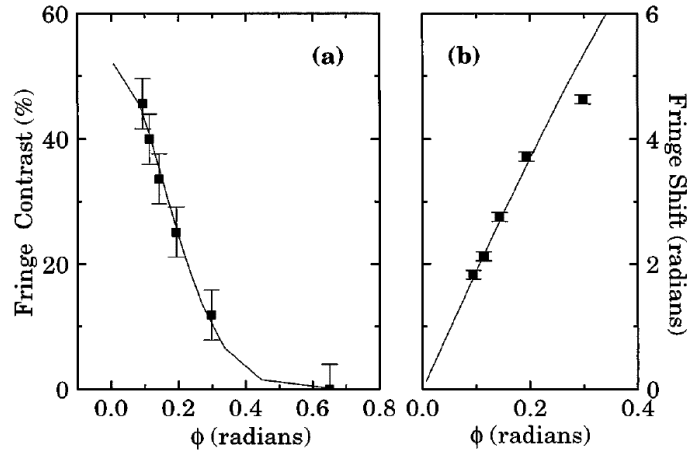
Figure 5.11: Ramsey fringe contrast (a), and shift (b), from [9].

The two graphs in Figure 5.11 show the results of theoretical calculations (smooth curves) and of the actual measurements (data points with error bars) for the Ramsey fringe contrast (a), and for the fringe shift (b) both in function of the cavity field phase shift ϕ. The agreement between theory and measurement is indeed striking, and it is here that the Haroche–Ramsey experiment is so remarkable.

The almost total loss of contrast in Figure 5.10 (d) can be understood quite easily by invoking the Bloch sphere picture. Let us identify $\mid e \rangle$ with $\mid \uparrow \rangle$, and $\mid g \rangle$ with $\mid \downarrow \rangle$.

Just before the atom arrives at R_1 it is in the $\mid \uparrow \rangle$ state. The pulse in R_1 rotates it to $\mid \rightarrow \rangle$ or some other "equatorial" state. Then the atom enters the "cat" cavity and becomes entangled with the photons in it. The combined superposition state that emerges from this interaction

$$\frac{1}{\sqrt{2}} \left(\mid \uparrow \rangle \mid \alpha e^{i\phi} \rangle + \mid \downarrow \rangle \mid \alpha e^{-i\phi} \rangle \right)$$

is pure, but if we were to trace the photons out, we would find that the qubit itself is no longer in a pure state. We would find that it has *dephased*, as we have seen in Section 5.11.3. The dephasing parameter Γ in equation (5.28) would be quite large on account of the detuning being relatively small. Consequently, the exponent $e^{-\Gamma t}$ would kill r^x and r^y of the qubit, leaving only r^z. But an "equatorial" state of a qubit does not have any r^z. So the qubit emerges from the cavity C in a completely chaotic state $\boldsymbol{r} = 0$. The application of the next Rabi pulse in R_2 does nothing to this state, it remains chaotic, and when it is finally measured by D_g and D_e it returns a flat curve, as seen in Figure 5.11 (d).

The act of measurement commited by the "cat" destroys the original state of the qubit, but the information about it survives in the "cat's" ϕ. What more, tracing the qubit out from the combined state has much less effect on the "cat" than tracing the "cat" out of the combined state had on the qubit, because there is more of the "cat". And so, the "cat" will remain in the superposition

$$\frac{1}{\sqrt{2}} \left(\mid \alpha e^{i\phi} \rangle + \mid \alpha e^{-i\phi} \rangle \right)$$

or in a slightly mixed state that is very close to it.

This can be seen by sending another atom through C almost immediately after the first atom. The role of the second atom is to read the state of the field in the cavity.

This complicates the picture somewhat. The second atom entangles with the cavity field, adding or subtracting another ϕ to its phase. Furthermore, through the cavity field, the second atom also entangles with the first one. Were we to neglect the progressing dephasing of the cavity field, the combined state of both atoms and the cavity photons would be

$$\frac{1}{\sqrt{4}} \left(\mid e \rangle_2 \mid \alpha e^{i2\phi} \rangle \mid e \rangle_1 + \mid e \rangle_2 \mid \alpha \rangle \mid g \rangle_1 + \mid g \rangle_2 \mid \alpha \rangle \mid e \rangle_1 + \mid g \rangle_2 \mid \alpha e^{-i2\phi} \rangle \mid g \rangle_1 \right)$$

Tracing the cavity field out lets us evaluate probabilities of correlated detections for both atoms, P_{ee}, P_{eg}, P_{ge} and P_{gg}. These depend on the frequency of the field applied to the two Ramsey cavities R_1 and R_2, and on the time lapse between the two atoms, τ. But it turns out that the following combination of the probabilities

$$\eta = \frac{P_{ee}}{P_{ee} + P_{eg}} - \frac{P_{ge}}{P_{ge} + P_{gg}}$$

is largely independent of the Ramsey (R_1 and R_2) pulse frequency. This quantity is our measure of the cavity field coherence. It should be $1/2$ for very short times τ (it is actually less than this in Figure 5.12 because of the same experimental difficulties that reduce the expected contrast of Ramsey fringes) and we expect it to decay exponentially with τ.

Why should the "cat" itself dephase? It dephases because it is, in turn, entangled with the environment—for example its power source marked S in Figure 5.9.
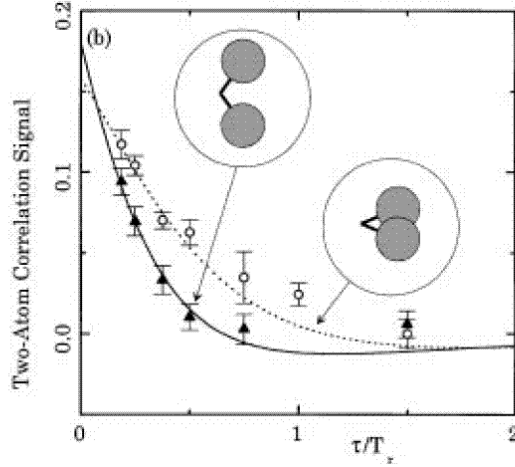
Figure 5.12: Two atom correlation signal η in function of τ/T_r, where τ is the delay between the two atoms and T_r is the cavity photon lifetime of $160\,\mu$s, from [9].

Figure 5.12 shows what happens for the two detunings of $\delta = 2\pi \times 170\,$kHz (circles and the dashed curve), and of $\delta = 2\pi \times 70\,$kHz (triangles and the smooth curve), with the cavity C filled by 3.3 photons on average. Circles and triangles correspond to the measured data and the curves correspond to theoretical predictions.

The field in the cavity indeed decoheres rapidly and on the time scale that is well short of the photon lifetime $T_r = 160\,\mu$s. Furthermore, the field configuration that corresponds to stronger entanglement and yields a larger phase shift ϕ (triangles and the smooth line) decays faster.

The "cat" may remain longer in a superposition state if the two components of the state do not differ sufficiently to see the difference. The more they do, the faster the dephasing of the state.

6 The Controlled NOT Gate

A controlled–NOT gate (CNOT for short) is a biqubit gate that makes quantum computing possible. In combination with single qubit unitary gates, the controlled–NOT gate is a universal gate of quantum computation, meaning that every unitary operation on an n-qubit register can be implemented by combining controlled–NOT and single qubit unitary gates. In this the gate is similar to a NAND gate known from classical computing. Every Boolean and arithmetic operation on an n-bit register can be implemented by a combination of NAND gates.

Although the definition of a controlle–NOT gate is strikingly simple, the gate is very hard to construct. The previous chapter explains why: we have to perform a controlled operation on a biqubit. This is hard. Biqubits are hard to make and hard to control, while maintaing their biqubitness, that is, entanglement, and the state's purity at the same time.

This chapter will illustrate this point by discussing various implementations and simple uses of a controlled–NOT gate.

But first things first. What is a controlled NOT gate?

It is a single qubit NOT gate that is controlled by another qubit.

Consider a biqubit $\mid \psi \rangle \otimes \mid \eta \rangle$. Let the first qubit, the one on the left, be the control qubit and the one on the right the object qubit of the gate. We define the gate by saying that when the control qubit is in the $\mid 0 \rangle$ state, then the gate is inactive. It leaves the object qubit as it is. But when the control qubit is in the $\mid 1 \rangle$ state, then the gate flips the object qubit:

$$
\begin{aligned}
\mid 0 \rangle \mid 0 \rangle &\rightarrow \mid 0 \rangle \mid 0 \rangle \\
\mid 0 \rangle \mid 1 \rangle &\rightarrow \mid 0 \rangle \mid 1 \rangle \\
\mid 1 \rangle \mid 0 \rangle &\rightarrow \mid 1 \rangle \mid 1 \rangle \\
\mid 1 \rangle \mid 1 \rangle &\rightarrow \mid 1 \rangle \mid 0 \rangle
\end{aligned}
\tag{6.1}
$$

It is convenient to replace the explicit tensor notation with the following decimal labeling of the biqubit states

$$
\begin{aligned}
\mid 0 \rangle \mid 0 \rangle &\equiv \mid \mathbf{0} \rangle \\
\mid 0 \rangle \mid 1 \rangle &\equiv \mid \mathbf{1} \rangle \\
\mid 1 \rangle \mid 0 \rangle &\equiv \mid \mathbf{2} \rangle \\
\mid 1 \rangle \mid 1 \rangle &\equiv \mid \mathbf{3} \rangle
\end{aligned}
\tag{6.2}
$$

Using this representation we can rewrite equations (6.1) in the decimal form

$$
\begin{aligned}
|\,\mathbf{0}\,\rangle &\rightarrow |\,\mathbf{0}\,\rangle \\
|\,\mathbf{1}\,\rangle &\rightarrow |\,\mathbf{1}\,\rangle \\
|\,\mathbf{2}\,\rangle &\rightarrow |\,\mathbf{3}\,\rangle \\
|\,\mathbf{3}\,\rangle &\rightarrow |\,\mathbf{2}\,\rangle
\end{aligned}
\tag{6.3}
$$

and this leads to the following matrix representation of the gate

$$
\mathrm{CNOT} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0
\end{pmatrix}
\tag{6.4}
$$

6.1   Physical Realization of The Controlled NOT Gate

6.1.1   An NMR Gate

6.1.2   An All-Optical Gate

6.1.3   A Calcium Ion Gate

6.2   Universality of The Controlled-NOT Gate

7 Simple Quantum Circuits

REPLACE this line with the first line of your chapter.

8 Open Questions

REPLACE this line with the first line of your chapter.

A Quaternions and Pauli Matrices

A.1 Hamilton quaternions

$$q = a + bi + cj + dk$$

where

$$ii = jj = kk = -1$$

and

$$ijk = -1$$

from which it follows that

$$ij = -ji = k$$
$$jk = -kj = i$$
$$ki = -ik = j$$

Extraction of quaternion components:

$$
\begin{aligned}
a &= \Re(q) \\
b &= \Im_i(q) = -\Re(iq) \\
c &= \Im_j(q) = -\Re(jq) \\
d &= \Im_k(q) = -\Re(kq)
\end{aligned}
$$

For

$$
\begin{aligned}
a &= a^x i + a^y j + a^z k \\
b &= b^x i + b^y j + b^z k
\end{aligned}
$$

the following holds

$$ab = -\vec{a} \cdot \vec{b} + \left(\vec{a} \times \vec{b}\right)^x i + \left(\vec{a} \times \vec{b}\right)^y j + \left(\vec{a} \times \vec{b}\right)^z k$$

A.2 Pauli quaternions

$$q = a + b\sigma_x + c\sigma_y + d\sigma_z$$

where

$$\begin{aligned}
\boldsymbol{\sigma}_x &= i\boldsymbol{i} \\
\boldsymbol{\sigma}_y &= i\boldsymbol{j} \\
\boldsymbol{\sigma}_z &= i\boldsymbol{k}
\end{aligned}$$

and

$$\boldsymbol{\sigma}_x\boldsymbol{\sigma}_x = \boldsymbol{\sigma}_y\boldsymbol{\sigma}_y = \boldsymbol{\sigma}_z\boldsymbol{\sigma}_z = 1$$

and

$$\boldsymbol{\sigma}_x\boldsymbol{\sigma}_y\boldsymbol{\sigma}_z = i$$

from which it follows that

$$\begin{aligned}
\boldsymbol{\sigma}_x\boldsymbol{\sigma}_y &= -\boldsymbol{\sigma}_y\boldsymbol{\sigma}_x = i\boldsymbol{\sigma}_z \\
\boldsymbol{\sigma}_y\boldsymbol{\sigma}_z &= -\boldsymbol{\sigma}_z\boldsymbol{\sigma}_y = i\boldsymbol{\sigma}_x \\
\boldsymbol{\sigma}_z\boldsymbol{\sigma}_x &= -\boldsymbol{\sigma}_x\boldsymbol{\sigma}_z = i\boldsymbol{\sigma}_y
\end{aligned}$$

These can be encapsulated into

$$\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j = \delta_{ij}\mathbf{1} + i\sum_k \epsilon_{ijk}\boldsymbol{\sigma}_k$$

Extraction of quaternion components:

$$\begin{aligned}
a &= \Re\left(\boldsymbol{q}\right) \\
b &= \Im_x\left(\boldsymbol{q}\right) = \Re(\boldsymbol{\sigma}_x\boldsymbol{q}) \\
c &= \Im_y\left(\boldsymbol{q}\right) = \Re(\boldsymbol{\sigma}_y\boldsymbol{q}) \\
d &= \Im_z\left(\boldsymbol{q}\right) = \Re(\boldsymbol{\sigma}_z\boldsymbol{q})
\end{aligned}$$

For

$$\begin{aligned}
\boldsymbol{a} &= a^x\boldsymbol{\sigma}_x + a^y\boldsymbol{\sigma}_y + a^z\boldsymbol{\sigma}_z \\
\boldsymbol{b} &= b^x\boldsymbol{\sigma}_x + b^y\boldsymbol{\sigma}_y + b^z\boldsymbol{\sigma}_z
\end{aligned}$$

the following holds

$$\boldsymbol{a}\boldsymbol{b} = \vec{a}\cdot\vec{b} + i\left(\left(\vec{a}\times\vec{b}\right)^x\boldsymbol{\sigma}_x + \left(\vec{a}\times\vec{b}\right)^y\boldsymbol{\sigma}_y + \left(\vec{a}\times\vec{b}\right)^z\boldsymbol{\sigma}_z\right)$$

A.3 Pauli matrices

$$q = a\mathbf{1} + b\boldsymbol{\sigma}_x + c\boldsymbol{\sigma}_y + d\boldsymbol{\sigma}_z$$

where

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and where $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ have all commutation properties identical with Pauli quaternions. Pauli matrices are a *representation* of Pauli quaternions.

For all Pauli quaternions/matrices

$$2\Re = \mathrm{Tr}$$

For

$$q = a\mathbf{1} + b\boldsymbol{\sigma}_x + c\boldsymbol{\sigma}_y + d\boldsymbol{\sigma}_z$$

$$a = \frac{1}{2}\mathrm{Tr}\,(q)$$

$$b = \frac{1}{2}\mathrm{Tr}\,(\boldsymbol{\sigma}_x q)$$

$$c = \frac{1}{2}\mathrm{Tr}\,(\boldsymbol{\sigma}_y q)$$

$$d = \frac{1}{2}\mathrm{Tr}\,(\boldsymbol{\sigma}_z q)$$

Canonical basis in the space of 2×2 matrices expressed in terms of Pauli matrices:

$$M_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\mathbf{1} + \boldsymbol{\sigma}_z)$$

$$M_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\boldsymbol{\sigma}_x + i\boldsymbol{\sigma}_y)$$

$$M_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}(\boldsymbol{\sigma}_x - i\boldsymbol{\sigma}_y)$$

$$M_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}(\mathbf{1} - \boldsymbol{\sigma}_z) \tag{A.1}$$

B   Bi-qubit Probability Matrices

$$\begin{aligned}
\varsigma_1 &= e_0 + e_1 + e_2 + e_3 \\
\varsigma_x &= e_2 \\
\varsigma_y &= e_3 \\
\varsigma_z &= e_0 - e_1
\end{aligned}$$

$$\begin{aligned}
\varsigma_{1A} \otimes \varsigma_{1B} &= (e_{0A} + e_{1A} + e_{2A} + e_{3A}) \otimes (e_{0B} + e_{1B} + e_{2B} + e_{3B}) \\
&= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \\
\varsigma_{xA} \otimes \varsigma_{1B} &= e_{2A} \otimes (e_{0B} + e_{1B} + e_{2B} + e_{3B}) \\
&= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
\varsigma_{1A} \otimes \varsigma_{xB} &= (e_{0A} + e_{1A} + e_{2A} + e_{3A}) \otimes e_{2B} \\
&= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\
\varsigma_{yA} \otimes \varsigma_{1B} &= e_{3A} \otimes (e_{0B} + e_{1B} + e_{2B} + e_{3B}) \\
&= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \\
\varsigma_{1A} \otimes \varsigma_{yB} &= (e_{0A} + e_{1A} + e_{2A} + e_{3A}) \otimes e_{3B} \\
&= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
\varsigma_{zA} \otimes \varsigma_{1B} &= (e_{0A} - e_{1A}) \otimes (e_{0B} + e_{1B} + e_{2B} + e_{3B})
\end{aligned}$$

$$
= \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

$$
\boldsymbol{\varsigma}_{1A} \otimes \boldsymbol{\varsigma}_{zB} = (\boldsymbol{e}_{0A} + \boldsymbol{e}_{1A} + \boldsymbol{e}_{2A} + \boldsymbol{e}_{3A}) \otimes (\boldsymbol{e}_{0B} - \boldsymbol{e}_{1B})
$$

$$
= \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}
$$

$$
\boldsymbol{\varsigma}_{xA} \otimes \boldsymbol{\varsigma}_{xB} = \boldsymbol{e}_{2A} \otimes \boldsymbol{e}_{2B}
$$

$$
= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

$$
\boldsymbol{\varsigma}_{yA} \otimes \boldsymbol{\varsigma}_{yB} = \boldsymbol{e}_{3A} \otimes \boldsymbol{e}_{3B}
$$

$$
= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

$$
\boldsymbol{\varsigma}_{zA} \otimes \boldsymbol{\varsigma}_{zB} = (\boldsymbol{e}_{0A} - \boldsymbol{e}_{1A}) \otimes (\boldsymbol{e}_{0B} - \boldsymbol{e}_{1B})
$$

$$
= \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

C Tensor Products of Pauli Matrices

$$\mathbf{1} \otimes \mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\equiv \begin{pmatrix} 1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & 0\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ 0\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & 1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (C.1)$$

$$\boldsymbol{\sigma}_x \otimes \boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\equiv \begin{pmatrix} 0\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 1\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ 1\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 0\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad (C.2)$$

$$\boldsymbol{\sigma}_y \otimes \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\equiv \begin{pmatrix} 0\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} & -i\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ i\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} & 0\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \qquad (C.3)$$

$$\boldsymbol{\sigma}_z \otimes \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$
\equiv
\begin{pmatrix}
1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & 0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & -1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\end{pmatrix}
$$

$$
=
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\tag{C.4}
$$

References

[1] *Encyklopedia Fizyki*, volume 2. Państwowe Wydawnictwo Naukowe, Warszawa, 1973.

[2] Markus Aspelmeyer, Hannes R. Böhm, Tsewang Gyatso, Thomas Jennewein, Rainer Kaltenbaek, Michael Lindenthal, Gabriel Molina-Terriza, Andreas Poppe, Kevin Resch, Michael Taraba, Rupert Ursin, Philip Walther, and Anton Zeilinger. Long-distance free-space distribution of quantum entanglement. *Science*, 301:621–623, August 2003.

[3] J. S. Bell. On the einstein podolsky rosen paradox. *Physics*, 1:195–200, 1964.

[4] J. S. Bell. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge University Press, Cambridge, 1987.

[5] A. J. Berkley, H. Xu, R. C. Ramos, M. A. Gubrud, F. W. Strauch, P. R. Johnson, J. R. Anderson, A. J. Dragt, C. J. Lobb, and F. C. Wellstood. Entangled macroscopic quantum states in two superconducting qubits. *Science*, 300:1548–1550, June 2003.

[6] M. V. Berry. Quantal phase factors accompanying adiabatic changes. *Proceedings of the Royal Society of London*, A 392:45–57, 1984.

[7] B. B. Blinov, D. L. Moehring, L.-M. Duan, and C. Monroe. Observation of entanglement between a single trapped atom and a single photon. *Nature*, 428:153–157, March 2004.

[8] D. Bohm and B. J. Hiley. *The Undivided Universe*. Routledge, London and New York, 1996.

[9] M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J. M. Raimond, and S. Haroche. Observing the progressive decoherence of the "meter" in a quantum measurement. *Physical Review Letters*, 77(24):4887–4890, December 1996.

[10] M. Brune, F. Schmidt-Kaler, A. Maali, J. Dreyer, E. Hagley, J. M. Raimond, and S. Haroche. Quantum rabi oscillation: A direct test of field quantization in a cavity. *Physical Review Letters*, 76(11):1800–1803, March 1996.

[11] J. F. Clauser, R. A. Hold, M. A. Horne, and A. Shimony. *Physical Review Letters*, 23:880, 1969.

[12] L. Davidovich, M. Brune, J. M. Raimond, and S. Haroche. Mesoscopic quantum coherences in cavity qed: Preparation and decoherence monitoring schemes. *Physical Review A*, 53(3):1295–1309, March 1996.

[13] B. S. DeWitt and N. Graham, editors. *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton University Press, Princeton, New Jersey, 1973.

[14] Thomas Durt and Yves Pierseaux. Bohm's interpretation and maximally entangled states. *Physical Review A*, 66:052109, 2002.

[15] A. Einstein. über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt. *Annalen der Physik*, 17:132–148, 1905.

[16] A. Einstein, B. Podolsky, , and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, 1935.

[17] Berthold-Georg Englert and Nasser Metwally. Remarks on 2-q-bit states. *arXiv:quant-ph*, (0007053v1), July 2000.

[18] A. Aspect et al. Experimental tests of realistic local theories via bell's theorem. *Physical Review Letters*, 47:460, 1981.

[19] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics*, volume 1. Addison-Wesley Publishing Company, July 1975.

[20] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics*, volume 3. Addison-Wesley Publishing Company, 1979.

[21] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics*, volume 2. Addison-Wesley Publishing Company, 1981.

[22] M. Fitting, editor. *Beyond Two: Theory and Applications of Multiple Valued Logic*. Physica Verlag, December 2002. ISBN 3790815411.

[23] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger. Bell's theorem without inequalities. *American Journal of Physics*, 58:1131–1143, 1990.

[24] D. M. Greenberger, M. A. Horne, and A. Zeilinger. *Bell's Theorem, Quantum Theory, and Conception of the Universe*, pages 73–76. Kluwer Academic, Dordrecht, 1989.

[25] Lucien Hardy. Quantum mechanics, local realistic theories, and lorentz-invariant realistic theories. *Physical Review Letters*, 68(20):2981–2984, May 1992.

[26] Lucien Hardy. Quantum theory from five reasonable axioms. *arXiv:quant-ph*, (0101012v4), September 2001.

[27] S. Haroche. Personal communication, August 2006.

[28] S. Haroche, M. Brune, and J. M. Raimond. Experiments with single atoms in a cavity: entanglement, schrödinger's cats and decoherence. *Philosophical Transactions of the Royal Society*, 355:2367, 1997.

[29] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223:1–8, November 1996.

[30] Paweł Horodecki. Separability criterion and inseparable mixed states with positive partial transposition. *Physics Letters A*, 232:333–339, August 1997.

[31] William T. M. Irvine, Juan F. Hodelin, Christoph Siomon, and Dirk Bouwmeester. Realization of hardy's thought experiment with photons. *Physical Review Letters*, 95:030401, July 2005.

[32] E. T. Jaynes and F. W. Cummings. *Proceedings of IEEE*, 51:89, 1963.

[33] Jonathan A. Jones, Vlatko Vedral, Artur Ekert, and Giuseppe Castagnoli. Geometric quantum computation using nuclear magnetic resonance. *Nature*, 403:869–871, February 2000.

[34] Brian Julsgaard, Alexander Kozhekin, and Eugene S. Polzik. Experimental long-lived entanglement of two macroscopic objects. *Nature*, 413:400–403, September 2001.

[35] *et al.* (140 authors) K. Hagiwara. Review of particle physics. *Physical Review D*, 66:010001 (974 pages), 2002.

[36] B. E. Kane. A silicon-based nuclear spin quantum computer. *Nature*, 393:133, 1998.

[37] B. E. Kane, N. S. McAlpine, A. S. Dzurak, R. G. Clark, G. J. Milburn, He Be Sun, and H. Wiseman. Single spin measurement using single electron transistors to probe two electron systems. *Physical Review B*, 61:2961, 2000.

[38] George J. Klir and Bo Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, May 1995. ISBN 0131011715.

[39] Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. W. H. Freeman and Company, 1973.

[40] G. Nogues, A. Rauschenbeutel, S. Osnaghi, M. Brune, J. M. Raymond, and S. Haroche. Seeing a single photon without destroying it. *Nature*, 400:239, 1999.

[41] A. K. Pati and A. K. Rajagopal. Inconsistencies of the adiabatic theorem and the berry phase. *arXiv:quant-ph*, (0405129v1), May 2004.

[42] Wolfgang Pauli. On the connexion between the completion of electron groups in an atom with the complex structure of spectra. *Zeitschrift für Physik*, 31:765, 1925.

[43] Roger Penrose. *The Road to Reality: A Complete Guide to the Laws of the Universe*. Knopf, February 2005. ISBN 0679454438.

[44] A. Peres and W. H. Zurek. Is quantum theory universally valid? *American Journal of Physics*, 50(9):807–810, September 1982.

[45] Asher Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers, 1995. Reprinted 1998, ISBN 0-7923-3632-1.

[46] Asher Peres. Separability criterion for density matrices. *Physical Review Letters*, 77(8):1413–1415, August 1996.

[47] J. M. Raimond and S. Haroche. Atoms and cavities: the birth of a schrödinger cat of the radiation field. In T. Asakura, editor, *International trends in optics and photonics*, pages 40–53. ICO IV, Springer Verlag, 1999.

[48] D. J. Richardson, A. I. Kilvington, K. Green, and S. K. Lamoreaux. Demonstration of berry's phase using stored ultracold neutrons. *Physical Review Letters*, 61(18):2030–2033, October 1988.

[49] Lee Smolin. *Three Roads to Quantum Gravity*. Perseus Books Group, 1 edition, August 2002. ISBN 0465078362.

[50] Matthias Steffen, M. Ansmann, Radoslaw C. Bialczak, N. Katz, Erik Lucero, R. McDermott, Matthew Neeley, E. M. Weig, A. N. Cleland, and John M. Martinis. Measurement of the entanglement of two superconducting qubits via state tomography. *Science*, 313:1423–1425, September 2006.

[51] Matthias Steffen, M. Ansmann, R. McDermott, N. Katz, Radoslaw C. Bialczak, Erik Lucero, Matthew Neeley, E. M. Weig, A. N. Cleland, and John M. Martinis. State tomography of capacitively shunted phase qubits with high fidelity. *Physical Review Letters*, 97(5):050502–1–3, August 2006.

[52] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, C. Urbina, D. Esteve, and M. H. Devoret. Manipulating the quantum state of an electrical circuit. *Science*, 296:886–889, May 2002.

[53] Stefan Weigert. Quantum time evolution in terms of nonredundant probabilities. *Physical Review Letters*, 84(5), January 2000.

[54] Jeng-Bang Yau, E. P. De Poortere, and M. Shayegan. Aharonov-bohm oscillations with spin: Evidence for berry's phase. *Physical Review Letters*, 88(14), April 2002.

Index